

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

100/124,176

12

EUROPEAN PATENT APPLICATION

21 Application number: **84303783.9**

51 Int. Cl.³: **C 12 N 15/00, C 12 P 21/00, A 61 K 37/36**

22 Date of filing: **05.06.84**

30 Priority: **06.06.83 US 501353**
20.06.83 US 506078

71 Applicant: **GENENTECH, INC., 460 Point San Bruno Boulevard, South San Francisco California 94080 (US)**

43 Date of publication of application: **19.12.84**
Bulletin 84/51

72 Inventor: **Lee, James Mon, c/o P.O. Box 5058, South San Francisco California 94080 (US)**
Inventor: **Ullrich, Axel, 433 Upper Terrace, San Francisco California 94117 (US)**

84 Designated Contracting States: **AT BE CH DE FR GB IT LI LU NL SE**

74 Representative: **Armitage, Ian Michael et al, MEWBURN ELLIS & CO. 2/3 Cursitor Street, London EC4A 1BQ (GB)**

54 **Human Insulin-like growth factor (IGF) and human epidermal growth factor (EGF) produced from a recombinant host, process, expression vector and recombinant host therefor, and IGF-containing pharmaceutical composition.**

57 **Human IGF and EGF are prepared by recombinant DNA technology.**

ALL IGF-1 Sequences Protected

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
205

Field of the Invention

This invention relates to the preparation of human IGF (insulin-like growth factor), in various forms, via recombinant DNA technology. Notably, the present invention provides for the preparation of human IGF as a mature protein product of expression, processing, and secretion in a recombinant DNA modified host organism. This invention thus provides for the production, isolation, and use of human IGF, in its various forms, as well as to the associated recombinant DNA technology by which it is prepared. In addition, the present invention relates to the similar preparation of a related protein, human EGF (Epidermal Growth Factor).

The present invention arises in part from the discovery of a novel system by which human IGF can be prepared by a recombinant host organism in the form of a discrete, mature protein. This is accomplished according to one aspect of the present invention by an expression system which permits the expression of the amino acid sequence of human IGF fused with at least a portion of the yeast alpha factor signal sequence, followed by processing of said signal sequence, and secretion of mature human IGF protein into the medium supporting the host organism. Thus, this novel aspect of the present invention, it is believed for the first time, permits the preparation, isolation, and utilization of human IGF as a discrete, mature protein. The present invention, in its broad compass, however, covers the preparation of the amino acid sequence of human IGF in other recombinant systems including bacteria and cell culture and includes, therefore, the expression of human IGF DNA sequences providing not only mature human IGF but also fusion product derivatives containing the amino acid sequence of IGF as the essential component. All such products have been found to be biologically active, hence useful as intended.

The publications and other materials hereof used to illuminate the background of the invention, and in particular cases, to provide additional details concerning its practice are incorporated herein by this reference and for convenience, are alphabetically and numerically referenced in the following text and respectively grouped in the appended bibliography.

10 Background of the Invention

A. Human IGF (Insulin-like Growth Factor)

Human IGF has been the subject of a fair amount of intensive study by past workers. A body of literature has been developed related to various aspects of this protein or series of proteins (see references A through L).

Insulin-like growth factors I and II have been isolated from human serum (A). The designation "insulin-like growth factor" or IGF was chosen to express the insulin-like effects and the insulin-like structure of these polypeptides which act as mitogens on a number of cells. The complete amino acid sequences of IGF-I and IGF-II have been determined (D,E). They are both single-chain polypeptides with three disulphide bridges and a sequence identity of 49 and 47 percent respectively, to human insulin A and B chains. The connecting peptide or C region is considerably shorter than the one of proinsulin and does not show any significant homology to it. (For a summary of earlier studies on the biological efforts of IGF, see Reference F).

IGF-I and IGF-II are growth promoting polypeptides occurring in human serum and human cerebral spinal fluid. Their structure is homologous to proinsulin. IGF-I seems to be produced by the liver along with a specific IGF-binding protein both of which are under

control of growth hormone. Thus, human IGF is considered to be an active growth promoting molecule that mediates the effect of human growth hormone.

5 It was perceived that the application of recombinant DNA and associated technologies would be a most effective way of providing the requisite large quantities of high quality human IGF for applied use to human beings as a growth factor. The goal was to produce human IGF either as biologically active fusion protein, or more
10 importantly, as a mature protein, as products of recombinant DNA technology from a host organism. Such materials would exhibit bioactivity admitting of their use clinically in the treatment of various growth affected conditions.

15 B. Recombinant DNA Technology

Recombinant DNA technology has reached the age of some sophistication. Molecular biologists are able to recombine various DNA sequences with some facility, creating new DNA entities capable
20 of producing copious amounts of exogenous protein product in transformed microbes and cell cultures. The general means and methods are in hand for the in vitro ligation of various blunt ended or "sticky" ended fragments of DNA, producing potent expression vehicles useful in transforming particular organisms, thus directing
25 their efficient synthesis of desired exogenous product. However, on an individual product basis, the pathway remains somewhat tortuous and the science has not advanced to a stage where regular predictions of success can be made. Indeed, those who portend
30 successful results without the underlying experimental basis, do so with considerable risk of inoperability.

DNA recombination of the essential elements, i.e., an origin of replication, one or more phenotypic selection characteristics, an expression promoter, heterologous gene insert and remainder vector,
35 generally is performed outside the host cell. The resulting

recombinant replicable expression vehicle, or plasmid, is introduced into cells by transformation and large quantities of the recombinant vehicle obtained by growing the transformant. Where the gene is properly inserted with reference to portions which govern the transcription and translation of the encoded DNA message, the resulting expression vehicle is useful to actually produce the polypeptide sequence for which the inserted gene codes, a process referred to as expression. The resulting product may be obtained by lysing, if necessary, the host cell, in microbial systems, and recovering the product by appropriate purification from other proteins.

In practice, the use of recombinant DNA technology can express entirely heterologous polypeptides--so-called direct expression--or alternatively may express a heterologous polypeptide fused to a portion of the amino acid sequence of a homologous polypeptide. In the latter cases, the intended bioactive product is sometimes rendered bioinactive within the fused, homologous/heterologous polypeptide until it is cleaved in an extracellular environment. See references (M) and (N).

Similarly, the art of cell or tissue cultures for studying genetics and cell physiology is well established. Means and methods are in hand for maintaining permanent cell lines, prepared by successive serial transfers from isolate normal cells. For use in research, such cell lines are maintained on a solid support in liquid medium, or by growth in suspension containing support nutriment. Scale-up for large preparations seems to pose only mechanical problems. For further background, attention is directed to references (O) and (P).

Likewise, protein biochemistry is a useful, indeed necessary, adjunct in biotechnology. Cells producing the desired protein also produce hundreds of other proteins, endogenous products of the cell's metabolism. These contaminating proteins, as well as other

compounds, if not removed from the desired protein, could prove toxic if administered to an animal or human in the course of therapeutic treatment with desired protein. Hence, the techniques of protein biochemistry come to bear, allowing the design of separation procedures suitable for the particular system under consideration and providing a homogeneous product safe for intended use. Protein biochemistry also proves the identity of the desired product, characterizing it and ensuring that the cells have produced it faithfully with no alterations or mutations. This branch of science is also involved in the design of bioassays, stability studies and other procedures necessary to apply before successful clinical studies and marketing can take place.

Summary of the Invention

The present invention is based upon the discovery that recombinant DNA technology can be used successfully to produce human IGF and related protein, human EGF, preferably in direct form and in amounts sufficient to initiate and conduct animal and clinical testing as prerequisites to market approval. The products human IGF and EGF are suitable for use in all of their forms as produced according to the present invention, viz. in the prophylactic or therapeutic treatment of human beings for various growth associated conditions or diseases. Accordingly, the present invention, in one important aspect, is directed to methods of treating growth conditions in human subjects using human IGF or human EGF, and suitable pharmaceutical compositions thereof, prepared in accordance with the methods and means of the present invention.

The present invention further comprises essentially pure, mature human IGF, as a product of expression, processing, and secretion in a recombinant host organism. Such human IGF is free from association with N-terminus amino acid sequence derivable from the expression systems that can be employed to prepare the material. Thus, while the present invention is directed to the preparation of

polypeptides comprising the amino acid sequence of IGF, a notable aspect of the present invention involves the production of the mature human IGF directly into the medium of the recombinant host organism employed. The present invention is also directed to
5 replicable DNA expression vehicles harboring gene sequences encoding human IGF and human EGF in expressible form, to microorganism strains or cell cultures transformed with such vehicles and to microbial or cell cultures of such transformants capable of producing amino acid sequences of human IGF and human EGF. In still
10 further aspects, the present invention is directed to various processes useful for preparing said genes sequences, DNA expression vehicles, microorganisms and cell cultures and specific embodiments thereof. Still further, this invention is directed to the preparation of fermentation cultures of said microorganisms and cell
15 cultures.

Brief Description of the Drawings

Figure 1 represents the chemically synthesized DNA strands used
20 in the construction of expression vectors for human IGF.

Figure 2 shows the completed double stranded DNA of Figure 1.

Figure 3 show the fragments of DNA of Figure 2 after restriction
25 by EcoRI and PstI and Bam HI.

Figure 4 depicts the ligation of parts 1 and 2 of Figure 3 into PBR322.

30 Figure 5 show parts 3 and 4 of IGF-I right half.

Figure 6 depicts the ligation of parts 3 and 4 of Figure 5 into the vector of Figure 4.

35 Figure 7 shows a sequence of DNA and deduced fusion protein containing IGF-I.

0390L/7

Figure 8 shows a sequence of DNA and deduced short fusion protein containing IGF-I.

Figure 9 depicts a plasmid used in the present construction.

Figure 10 shows the DNA and protein sequence of IGF-I fused with alpha factor pre-pro sequence.

Figure 11 is a vector containing alpha factor promotor and pre-pro sequence fused to IGF-I.

Figure 12 shows the yeast invertase signal fused to IGF-I.

Figure 13 shows the parental plasmid containing the yeast PGK promotor.

Figure 14 depicts a yeast expression vector containing PGK promotor, invertase signal and human IGF-I gene.

Figure 15 is the synthetic DNA used to construct the coding sequence of mature human EGF.

Figure 16 depicts the yeast alpha factor "pre-pro" sequence fused to the human EGF coding sequence.

Figure 17 depicts the yeast invertase signal sequence fused to the human EGF coding sequence.

Figure 18 shows the coding sequence for human IGF-II.

Detailed Description

A. Definitions

As used herein, "human IGF" and "human EGF" denotes human

insulin-like growth factor and human epidermal growth factor, produced by microbial or cell cultures systems and bioactive forms comprising the amino acid sequence corresponding to human IGF and human EGF otherwise native to human tissue. The human IGF and EGF proteins produced herein have been defined by means of DNA, gene, and deductive amino acid sequencing. It will be understood that inasmuch as natural allelic variations exist and occur from individual to individual, as demonstrated by (an) amino acid difference(s) in the overall sequence or by deletions, substitutions, insertions, inversions, or additions of one or more amino acids of said sequences, the present invention is intended to embrace of all such allelic variations of the two molecules involved. In addition, the location of and the degree of glycosylation depends upon the nature of the recombinant host organism employed and such variations as may occur are included within the ambit of this invention. Finally, the potential exists in the use of DNA technology for the preparation of various derivatives of human IGF and human EGF by simple modification of the underlying gene sequence for such molecules. Such modifications could be accomplished by means of site directed mutagenesis of the underlying DNA, as an example. All such modifications resulting in derivatives of human IGF and human EGF are included within the scope of the present invention so long as the essential characteristic human IGF and human EGF activities remain unaffected in kind.

"Essentially pure form" when used to describe the state of human IGF or human EGF produced by this invention means that the proteins are free of proteins or other materials normally associated with human IGF or human EGF when produced by non-recombinant cells, i.e. in their "native" environments.

"Expression vector" includes vectors which are capable of expressing DNA sequences contained therein, where such sequences are operably linked to other sequences capable of effecting their expression, i.e., promotor/operator sequences. In sum, "expression

vector" is given a functional definition: any DNA sequence which is capable of effecting expression of a specified DNA code disposed therein. In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer to
5 circular double stranded DNA loops which in their vector form, are not bound to the chromosome. In the present specification, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which
10 function equivalently and which become known in the art subsequently.

"Recombinant host cells" refers to cells which have been transformed with such vectors. Thus, the human IGF and human EGF
15 molecules produced by such cells can be referred to as "recombinant human IGF" and "recombinant human EGF".

B. Host Cell Cultures and Vectors

20 The vectors and methods disclosed herein are suitable for use in host cells over a wide range of prokaryotic and eukaryotic organisms.

In general, of course, prokaryotes are preferred for cloning of DNA sequences in constructing the vectors useful in the invention.
25 For example, E. coli K12 strain 294 (ATCC No. 31446) is particularly useful. Other microbial strains which may be used include E. coli strains such as E. coli B, and E. coli X1776 (ATCC No. 31537). The aforementioned strains, as well as E. coli W3110 (F⁻, λ ⁻, prototrophic, ATCC No. 27325), bacilli such as Bacillus subtilis,
30 and other enterobacteriaceae such as Salmonella typhimurium or Serratia marcesans, and various pseudomonas species may be used. These examples are, of course, intended to be illustrative rather than limiting.

35 In general, plasmid vectors containing replicon and control

sequences which are derived from species compatible with the host cell are used in connection with these hosts. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, E. coli is typically transformed using pBR 322, a plasmid derived from an E. coli species (Bolivar, et al., Gene 2: 95 (1977)). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own proteins. Those promoters most commonly used in recombinant DNA construction include the β -lactamase (penicillinase) and lactose promoter systems (Chang et al, Nature, 275: 615 (1978); Itakura, et al, Science, 198: 1056 (1977); (Goeddel, et al Nature 281: 544 (1979)) and a tryptophan (trp) promoter system (Goeddel, et al, Nucleic Acids Res., 8: 4057 (1980); EPO Appl Publ No. 0036776). While these are the most commonly used, other microbial promoters have been discovered and utilized, and details concerning their nucleotide sequences have been published, enabling a skilled worker to ligate them functionally with plasmid vectors (Siebenlist, et al, Cell 20: 269 (1980)).

In addition to prokaryotes, eukaryotic microbes, such as yeast cultures may also be used. Saccharomyces cerevisiae, or common baker's yeast is the most commonly used among eukaryotic microorganisms, although a number of other strains are commonly available. For expression in Saccharomyces, the plasmid YRp7, for example, (Stinchcomb, et al, Nature, 282: 39 (1979); Kingsman et al, Gene, 7: 141 (1979); Tschemper, et al, Gene, 10: 157 (1980)) is commonly used. This plasmid already contains the trp1 gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC No. 44076 or PEP4-1 (Jones, Genetics, 85: 12 (1977)). The presence of the trp1 lesion as a characteristic of the yeast host cell genome then provides an

effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoting sequences in yeast vectors include the promoters for 3-phosphoglycerate kinase (Hitzeman, et al., J. Biol. Chem., 255: 2073 (1980)) or other glycolytic enzymes (Hess, et al, J. Adv. Enzyme Reg., 7: 149 (1963); Holland, et al, Biochemistry, 17: 4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. In constructing suitable expression plasmids, the termination sequences associated with these genes are also ligated into the expression vector 3' of the sequence desired to be expressed to provide polyadenylation of the mRNA and termination. Other promoters, which have the additional advantage of transcription controlled by growth conditions are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and

galactose utilization (Holland, *ibid.*). Any plasmid vector containing yeast-compatible promoter, origin of replication and termination sequences is suitable.

- 5 In addition to microorganisms, cultures of cells derived from multicellular organisms may also be used as hosts. In principle, any such cell culture is workable, whether from vertebrate or invertebrate culture. However interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure in recent years [10 Tissue Culture, Academic Press, Kruse and Patterson, editors (1973)]. Examples of such useful host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, and W138, BHK, COS-7 and MDCK cell lines. Expression vectors for such cells ordinarily 15 include (if necessary) an origin of replication, a promoter located in front of the gene to be expressed, along with any necessary ribosome binding sites, RNA splice sites, polyadenylation site, and transcriptional terminator sequences.
- 20 For use in mammalian cells, the control functions on the expression vectors are often provided by viral material. For example, commonly used promoters are derived from polyoma, Adenovirus 2, and most frequently Simian Virus 40 (SV40). The early and late promoters of SV40 virus are particularly useful because 25 both are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication (Fiers, et al, Nature, 273: 113 (1978) incorporated herein by reference. Smaller or larger SV40 fragments may also be used, provided there is included the approximately 250 bp sequence extending from the Hind III site 30 toward the Bgl I site located in the viral origin of replication. Further, it is also possible, and often desirable, to utilize promoter or control sequences normally associated with the desired gene sequence, provide such control sequences are compatible with 35 the host cell systems.

An origin of replication may be provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral (e.g. Polyoma, Adeno, VSV, BPV, etc.) source, or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter is often sufficient.

C. Methods Employed

If cells without formidable cell wall barriers are used as host cells, transfection is carried out by the calcium phosphate precipitation method as described by Graham and Van der Eb, Virology, 52: 546 (1978). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used.

If prokaryotic cells or cells which contain substantial cell wall constructions are used, the preferred method of transfection is calcium treatment using calcium chloride as described by Cohen, F.N. et al Proc. Natl. Acad. Sci. (USA), 69: 2110 (1972).

Construction of suitable vectors containing the desired coding and control sequences employ standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required.

Cleavage is performed by treating with restriction enzyme (or enzymes) in suitable buffer. In general, about 1 µg plasmid or DNA fragments is used with about 1 unit of enzyme in about 20 µl of buffer solution. (Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer.) Incubation times of about 1 hour at 37°C are workable. After incubations, protein is removed by extraction with phenol and chloroform, and the nucleic acid is recovered from the aqueous fraction by precipitation with ethanol.

If blunt ends are required, the preparation is treated for 15 minutes at 15° with 10 units of Polymerase I (Klenow), phenol-chloroform extracted, and ethanol precipitated.

5 Size separation of the cleaved fragments is performed using 6 percent polyacrylamide gel described by Goeddel, D., et al, Nucleic Acids Res., 8: 4057 (1980) incorporated herein by reference.

10 For ligation approximately equimolar amounts of the desired components, suitably end tailored to provide correct matching are treated with about 10 units T4 DNA ligase per 0.5 µg DNA. (When cleaved vectors are used as components, it may be useful to prevent religation of the cleaved vector by pretreatment with bacterial alkaline phosphatase.)

15 For analysis to confirm correct sequences in plasmids constructed, the ligation mixtures are used to transform E. coli K12 strain 294 (ATCC 31446), and successful transformants selected by ampicillin resistance where appropriate. Plasmids from the
20 transformants are prepared, analyzed by restriction and/or sequenced by the method of Messing, et al, Nucleic Acids Res., 9:309 (1981) or by the method of Maxam, et al, Methods in Enzymology, 65:499 (1980).

Examples

25 The following examples are intended to illustrate but not to limit the present invention.

Synthesis and Expression of Human IGF-1

30 Enzymes were obtained from the following suppliers:

 New England Biolabs: restriction enzymes, T4 DNA ligase

 Bethesda Research Labs: restriction enzymes, Bact. Alkaline

Phos.

Boehringer-Mannheim: E. coli DNA Polymerase I (Klenow)

P+L Biochemicals: Polynucleotide kinase, Terminal
Nucleotidyl Transferase

5

New England Nuclear:

pBR322 [oligo(dG)-tailed] DNA

Reagents:

10

BioRad: Bis Acrylamide, Acrylamide, TEMED

Sigma: Ammonium Persulfate

Amersham: 10218 γ ^{32}P ATP >5000 Ci/mmol; 10165 α ^{32}P dCTP
>400 Ci/mmol.

15

Solutions and Media: 1X TBE: .54M Tris Base, 0.54 M Boric
Acid, .017 M Na_2EDTA .

Difco: Yeast Nitrogenous Base (YNB); Tryptone, Yeast
Extract; Bacto-Agar; Casamino Acids.

Autoradiography:

20

Kodak X-O mat AR XAR-2 Film

Glass Beads:

0.45-0.50 mm B. Braun Melsungen AG

25

LB medium (per liter):

10g NaCl; 5g yeast extract; 10g tryptone; .17ml NaOH (50
percent)

LB Agar (per liter):

30

10g tryptone; 5g yeast extract; .5g NaCl; 15g Bacto-Agar;
adjusted to pH 7.5 with NaOH.

Antibiotics:

35

Tetracycline (5 $\mu\text{g/ml}$) in all mediums; Ampicillin (20 $\mu\text{g/ml}$)
in all mediums (plates or liquid)

M9 Medium (per liter):

6g Na_2HPO_4 (anhydrous); 1g NH_4Cl ; 3g KH_2PO_4 ; .5g
NaCl; 1 mM MgSO_4 ; 0.5 percent (w/v) glucose; 0.5 percent
(w/v) Casamino Acids; .0001 percent Thiamine-HCl.

5

YNB-CAA (per liter):

6.7g Yeast Nitrogenous Base (without Amino Acids); 10 mg
adenine; 10 mg uracil; 5g Casamino Acids; 20g Glucose.

10

YNB-CAA agar plates (per liter):

Same as YNB-CAA + 30g agar.

Standard Ligation Conditions:

15

10-fold molar excess of insert (or linker) to vector. 1X T4
DNA ligase buffer and 400-800 U T4 DNA ligase; 14° - 12-16
hours.

Standard Kination Conditions:

20

1X Polynucleotide kinase buffer; 15 U polynucleotide kinase;
37° 60 minutes; followed by reaction termination by heating
to 65° for 10 minutes.

1X Kinase Buffer:

25

70 mM Tris-HCl (pH 7.6); 10 mM MgCl_2 ; 5 mM DTT

1X T4 DNA Ligase Buffer:

50 mM Tris-HCl (pH 7.8); 10 mM MgCl_2 ; 20 mM DTT; 1 mM rATP.

Construction, Strategy and Selection of a DNA Sequence.

30

35

The 1° protein structure of the human IGF-1 molecule has been
determined (1). Based upon this protein sequence and the genetic
code, a DNA sequence coding for mature human IGF-1 protein,
including all possible base substitutions at any one base position,
was determined by computer analysis (Genentech Untrans Program).

Using a restriction site analysis program (Genentech Asearch Program), all potential restriction sites located in all possible DNA sequences consistently coding for the same protein were found. Three sites internal to the coding sequence were selected: PstI, BamHI, and AvalI. Two additional sites were placed at the ends, just outside of the coding sequence of the mature protein: one EcoRI site before the initiation codon, AUG, and the SalI site following the termination codon, TAG of the coding sequence. The choice of these sites facilitated the cloning of the coding sequence in separate parts, each of which subsequently could be excised and then assembled to form an intact synthetic IGF-1 gene. This construction involved the assembly of 4 parts, 2 parts forming the left half, 2 parts forming the right half. Each part consisted of two single strands of chemically synthesized DNA (see Fig. 1). Proposed synthetic fragments were also analyzed for internal complementarity.

The constructions used to generate these four parts employed the use of DNA Polymerase I repair synthesis of synthetic oligonucleotide substrates having 9-10 bp stretches of complementary sequence at their 3' termini. In the presence of DNA Polymerase I (Klenow) and the four deoxynucleoside triphosphates, these primer-templates were extended to become full-length double-stranded DNAs. To prevent priming at locations other than the desired portions as well as self-hybridizations, each set of single-stranded DNAs were analyzed by a computer program (Genentech Homology Program), and wherever possible, sequences which would have potentially led to hairpin loops, self-priming, or mis-priming, were eliminated by alternate codon usage. Each of these four double-stranded DNAs were synthesized to include 9-12 additional bp of non-IGF-1 coding DNA at each end (see Fig. 2). This additional DNA was included to allow generation of sticky ends by restriction enzyme digestion. The sticky ends thus formed facilitated the ligation of the double-stranded pieces to contiguous coding sections of the synthetic gene or into a cloning vehicle.

The 9-12 extra bp of double stranded DNA beyond the restriction site at the end of each part (see Fig. 2) allowed for the TdT-mediated formation of single-stranded oligodeoxycytidine strands at the 3' ends of each double-stranded DNA section. These oligodeoxycytidine tailed double-stranded DNAs could then be annealed into a complementary oligodeoxyguanosine tailed PstI site of a cloning vehicle. Once cloned, and sequenced to ensure the correct base sequences, the parts could be easily isolated and ligated following restriction enzyme cleavage at the restriction sites selected at the ends of each of the four parts, to form the intact synthetic IGF-1 gene.

The method used successfully here was similar to that described by Rossi *et al.* (28); however, attempts at the construction and cloning of the IGF-1 coding sequence using the Rossi *et al.* method (28) with only two base pairs of extra DNA beyond the restriction enzyme recognition sites repeatedly failed. The method employed here also differs from the Rossi *et al.* procedure (28) in that restriction sites placed at both ends of a double stranded DNA allow for the convenience of cloning each double stranded DNA fragment, individually, by (dC)-tailing and annealing into a (dG)-tailed vector, a method which in practice requires less of the double stranded DNA than three-part ligations.

Chemical Synthesis

Eight fragments, 43, 43, 46, 46, 46, 46, 54, and 46 bases in length (see Fig. 1), were chemically synthesized according to the method of Crea and Horn (2), the only change being the use of mesitylene nitrotriazole as the condensing agent rather than 2,4,6-triisopropyl benzenesulfonylchloride tetrazole.

The syntheses of the fragments were accomplished from the appropriate solid support (cellulose) by sequential addition of the appropriate fully-protected dimer- or trimer-blocks. The cycles

were carried out under the same conditions as described in the synthesis of oligothymidilic acid (see Crea *et al.*, *supra*). The final polymers were treated with base (aqueous conc. NH_3) and acid (80 percent HOAc), the polymer pelleted off, and the supernatant
5 evaporated to dryness. The residue, dissolved in 4 percent aq. NH_3 , was washed with ethyl ester (3X) and used for the isolation of the fully deprotected fragment.

Purification was accomplished by electrophoresis using 20
10 percent polyacrylamide gels. The pure oligonucleotide was ethanol precipitated following gel elution.

225-285 pmoles of each chemically synthesized fragment was mixed with an equivalent amount of the complementary single-stranded DNA
15 fragment (i.e. 1L+3L; 2L+4L; 1R+3R, 2R+4R) in the presence of deoxyribonucleoside triphosphates at a final concentration of 200 μM with the exception of dCTP. dCTP was added to a concentration of 5 μM as a α - ^{32}P -labeled isotope with a specific activity of 1000-2000 Ci/mmol) to allow easy monitoring of the repair-synthesis reaction
20 product. The reactions were carried out in a buffer containing a final concentration of 50mM Tris HCl pH 7.5; 20 mM MgCl_2 ; 20 mM DTT and 154 DNA Polymerase I (Klenow) in a reaction volume of 200 μl . Reactions were allowed to proceed at 4° for 12-18 hrs.

25 Upon completion, EDTA was added to a concentration of 25 mM. Sample buffer containing the mixes were phenol extracted, CHCl_3 extracted 2X, and products were EtOH precipitated. Pellets were taken up in .3 M NaOAc and the DNA reprecipitated with EtOH . After
30 dissolving the pellets in H_2O , the 1L+3L and 2L+4L products were then digested separately with PstI in 100 μl reaction mixes containing 1X PstI buffer (50mM $(\text{NH}_4)_2\text{SO}_4$, 20 mM Tris HCl pH 7.5, 10 mM MgCl_2), and 70 U PstI. After 4 hrs, EDTA was added to a concentration of 10 mM, and the material was ethanol
35 precipitated. Pellets were then taken up in .3 M NaOAc and reprecipitated, then taken up in H_2O . The PstI-digested 1L+3L

product was digested with EcoRI at 37° in a 100 µl reaction mix 1X EcoRI buffer (150 mM NaCl, 6 mM Tris HCl pH 7.5, 6 mM MgCl₂) and 70 U EcoRI. The PstI digested 2L+4L product was digested at 37° with BamHI in a 100 µl reaction mix in 1X BamHI Buffer (150 mM NaCl, 6 mM Tris HCl pH 7.9, 6 mM MgCl₂) and 70 U BamHI. After 4 hrs, EDTA was added to both mixtures, and sample buffer was added. They were electrophoresed on a 6 percent polyacrylamide slab gel. Six percent slab gels were cast with a mixture containing 6 percent (w/v) acrylamide (20 to 1 ratio of acrylamide to Bis acrylamide) 1X TBE, 1 percent APS and 0.1 percent TEMED. Reaction products were located on the gel by autoradiography and the band corresponding to the 45 bp EcoRI-PstI digested 1L+3L product (Part 1) (see Fig. 3) and the band corresponding to the 50 bp PstI-BamHI digested 2L+4L product (Part 2) (see Fig. 3) were excised from the gel, the material electroeluted in 0.2X TBE, phenol extracted, CHCl₃ extracted, and ethanol precipitated. Parts 1 and 2 were dissolved in H₂O.

Cloning Vector Prep.

Cloning vector was prepared by digesting 20 µg pBR322 (15) with 50 U EcoRI and 50 U BamHI, in 1X RI Buffer at 37° for 6 hr. After addition of EDTA to a concentration of 10 mM, sample buffer was added, and the mixture was run on a 5 percent polyacrylamide gel. The gel was developed by staining 10' in H₂O containing 5 µg/ml Et. Bromide, rinsing 2X in H₂O and placing upon a UV transilluminator (302 nm). the band corresponding to the ca. 3712 bp EcoRI-BamHI digested pBR322 molecules was cut from the gel. The DNA was electroeluted from the gel slice, phenol extracted, CHCl₃ extracted 2X, and ethanol precipitated. The pellet was dissolved in H₂O and was ready for ligation.

Ligation

In a three-part ligation (see Fig. 4), in which the molar ratio

of inserts to vector in the ligation reaction was approximately 10 to 1, parts 1 and 2 were ligated into the EcoRI-BamHI digested 322 vector in 1X T4 DNA ligase buffer (cont. 50 mM Tris HCl pH 7.8; 10 mM MgCl₂, 20 mM DTT, 1 mM rATP) and ~800 U T4 DNA ligase (NEB).

The reaction was carried out at 14° for 12-16 hrs.

Transformations

E. coli strain 294 was used as the transformation host, using the procedure of M. Dagert and S.D. Ehrlich (3). The transformed cells were plated on LB-agar plates containing ampicillin (20 µg/ml; LB-Amp-plates) and transformants were screened and grown in LB medium containing ampicillin at 20 µg/ml ampicillin. Transformants were screened using a modification of the rapid miniscreen method of Birnboim and Doly (4). Miniprep DNA prepared as such was digested with EcoRI and BamHI and run on polyacrylamide slab gels. Several transformants which illustrated a ca. 218 bp EcoRI-BamHI insert were grown in large scale and plasmids from each were isolated and sequenced according to the procedure of Maxam and Gilbert (5) to confirm the correct chemical synthesis and construction. The pBR322 vector containing the complete correct left half sequence of IGF-1 was called IGF-1 LH 322 (see Fig. 5).

Cloning of Fragments of the Right Half of IGF-1.

Using the identical conditions of DNA Polymerase I-mediated repair-synthesis, the two pairs of fragments comprising the right half of the synthetic IGF-1 were converted into double-stranded DNAs. After the DNA Polymerase I reactions, and without enzymatic digestion, the 1R+3R (Part III) and 2R+4R (Part IV) reactions were run on a 6 percent polyacrylamide slab gels. The 83 bp (Part III) and 91 bp (Part IV) bands were located by autoradiography and cut from the gel. After electroelution the ethanol precipitated double-stranded DNAs were dC-tailed (see Fig. 6) using the procedures of Villa-Komaroff et al. (6) and Rowenkamp and Firtel

(7). Reactions were carried out in 50 μ l vols. of 1X tailing mix (cont. .2M Pot. Cacodylate, 25 mM Tris HCl pH 5.9, 2 mM DTT, .5 mM CoCl_2) and 22 μ M dCTP. After prewarming at 37° for 10', the 150 second reaction was begun by the addition of 10-20 units of terminal
5 nucleotidyl transferase and terminated by addition of EDTA followed by phenol extraction, CHCl_3 extraction 2X, and ethanol precipitation.

These oligo (dC) tailed Parts III and IV were then separately
10 mixed with equimolar amounts of oligo (dG)-tailed PstI cut pBR322 vector in 50 μ l of 1X annealing buffer (.1M NaCl; 10 mM Tris HCl pH 7.8, 1 mM EDTA) at a final DNA concentration of 1-2 μ g/ml. After heating to 75°C, the mixes were gradually cooled to 4° over a period of 16 hr and the mix transformed into competent *E. coli* 294 cells
15 prepared according to the procedure of Dagert and Ehrlich (3). Transformed cells were plated on LB-Tetracycline-Agar plates and grown in LB-Tetracycline medium at tetracycline concentrations of 5 μ g/ml. Tetracycline resistant transformants were picked and plated onto LB-Ampicillin-Agar plates to check for insertions at the PstI
20 site. Several tetracycline resistant, Ampicillin-sensitive colonies for each Part 3 and 4 were miniscreened and those exhibiting insertions at the PstI locus were grown in large scale and sequenced by the Maxam and Gilbert technique (5) to confirm the correct DNA sequences of Parts 3 and 4.

25 Construction of an Intact Synthetic HuIGF-1 Coding Sequence

Preparation: Parts 3 and 4.

30 Parts 3 and 4 were separately removed from their vectors by digestions of 20 μ g of each vector with AvaII in 1X AvaII buffer (60 mM NaCl, 6 mM Tris-HCl (pH 8.0); 10 mM MgCl_2 ; 6 mM 2-mercaptoethanol) and 30 U of AvaII. After 6 hr., at 37°, EDTA was
35 added to the 150 μ l reactions to a concentration of 15 mM and the material phenol extracted, CHCl_3 extracted 2X and ethanol

precipitated. The Part 3 pellet was then taken up in 1X BamHI buffer and digested in a volume of 150 μ l with 30 U BamHI at 37° for 4 hr. The pellet containing Part 4 was digested with 30 U SalI in 150 μ l of 1X SalI buffer at 37° for 4 hr.

Both digests were then run on 6 percent polyacrylamide slab gels and stained. The 51 bp band representing Part 3 and the 62 bp band representing Part 4 were removed from the gels and the DNA electroeluted, phenol extracted, CHCl₃ extracted 2X and ethanol precipitated. Pellets were then taken up in H₂O and were ready for ligation.

Vector Preparation

20 μ g of the IGF-1 LH 322 vector was digested with 50 U of BamHI and 50 U of SalI in a 200 μ l reaction containing 1X BamHI buffer at 37° for 6 hr. After addition of EDTA to a concentration of 15 mM, the digestion mix was run on a 6 percent polyacrylamide slab gel, ethidium bromide stained and the 3814 bp band excised from the gel.

After electroelution, phenol extraction, chloroform extraction and ethanol precipitation, the DNA pellet was taken up in H₂O and was ready for ligation with Parts 3 and 4 in a three-part ligation. The ligation was performed under conditions described above for a three-part ligation (see Fig. 7). Parts 3 and 4 were present in the ligation mix at a 10-fold molar excess of inserts to vector. The mix was transformed into competent *E. coli* 294 cells prepared according to the Dagert and Ehrlich procedure (3) and plated onto LB-Ampicillin plates. Several transformants were miniscreened and two clones exhibiting a ca. 115 bp BamHI-SalI fragment were grown in large scale and their plasmids prepared. Both strands of the intact synthetic gene were sequenced by the Maxam-Gilbert technique (5) to confirm the correct sequence. The pBR322 plasmid containing the complete correct sequence coding for Human IGF-1 was called pBR322 HuIGF-1.

Human IGF-1 ExpressionIGF-1 Fusion Expression in Bacteria

Initial attempts were to obtain expression of IGF-1 as a fusion protein. To accomplish this, both the pNCV (9) and the pNCVsLE (10) expression vectors were used. (The pNCVsLE expression vector is a derivative of the pNCV vector and was prepared as follows: pNCV was treated with BgII, which cleaves at the 13 codon of the LE fusion. The site was converted to an EcoRI cleavage site using synthetic DNA, to give the expression vector pNCVsLE. The synthetic DNA introduced into the plasmid has the sequence:

5'-GATCCAGAATTC
5' GATCGAATTCTG and this sequence was introduced
into the plasmid:
GATCCAGAATTC
GTCTTAAGCTAG

As a strategy to release the fused human IGF-1 protein from the trp fusion protein, a linker was designed such that an enzymatic proteolysis method reported by Wunsch et al. (8) could be applied to this expression system. To accomplish this, a DNA linker:

ProAla
5'-AATTCCTGCCG -3'
3' GGGACGGCCAG-5'

was chemically synthesized by standard methods (2) which when linked to the trp fusion protein and the IGF-1 gene, coded for the amino acid residues Proline and Alanine followed by Glycine and Proline which are the first two amino acid residues of IGF-1 and preceded by Proline and Alanine together comprise a recognition site for a collagenase isolated from Clostridium histolyticum (11,12). This enzyme reportedly acts at such a site to cleave the alanine-glycine peptide bond.

To construct a DNA sequence coding for a fusion protein with a collagenase cleavage site, 30 µg pBR322 HuIGF-1 plasmid was cleaved with 50 U BamHI and 50 U PvuI enzyme in 200 µl 1X BamHI buffer at 37° for 6 hours. After addition of EDTA to a concentration of 15 mM, the reaction

mix was chromatographed on a 6 percent polyacrylamide slab gel. The smaller PvuI-BamHI fragment (~725 bp) was isolated and digested with 40 U AvaII in 150 μ l 1X Sau96I buffer (60 mM NaCl, 6 mM Tris-HCl pH 7.4, 15 mM MgCl₂, 6 mM 2-mercaptoethanol). After addition of EDTA to a concentration of 15 mM, the resulting mix chromatographed on a 6 percent polyacrylamide slab gel. The smaller Sau96I-BamHI fragment (~86 bp) was extracted from the gel, phenol extracted, chloroform extracted 2X, and ethanol precipitated. This fragment was ready for ligation.

200 pmols of linker fragments were kinased with 100 U polynucleotide kinase in 20 μ l of 1X polynucleotide kinase buffer (70 mM Tris-HCl (pH 7.6); 10 mM MgCl₂; 5 mM DTT; 1 mM rATP) at 37° for 1 hour. The reaction was terminated by heating to 55°C for 5 minutes. 100 pmols of the kinased linker fragments were ligated to the 86 bp Sau96I-BamHI fragment with 400 U of T4 DNA ligase in 30 μ l of 1X T4 DNA ligase buffer at 14° for 12-16 hours. The ligation reaction was terminated by addition of EDTA to a concentration of 15 mM followed by phenol extraction, chloroform extraction 2X, and ethanol precipitation. The pellet was then taken up in 1X BamHI buffer and digested in a 100 μ l reaction with 50 U of EcoRI and 50 U of BamHI at 37° for 6 hrs. After terminating the digestion with EDTA, the mixture was chromatographed on a 6 percent polyacrylamide slab gel and the newly created (~97 bp) EcoRI-BamHI fragment was extracted from the gel, and prepared for ligation. The vector to receive this new fragment was prepared by digesting 30 μ g pBR322 HuIGF-1 with 100 U of each EcoRI and BamHI in 200 μ l of 1X BamHI buffer at 37° for 8 hr. The reaction was terminated, chromatographed on a 6 percent polyacrylamide slab gel and the larger band (~3830 bp) representing the EcoRI-BamHI digested plasmid was isolated and the plasmid DNA extracted and prepared for ligation as above. In a 30 μ l ligation reaction containing a 10-fold molar excess of insert fragment to vector, the EcoRI-BamHI fragment was ligated into the EcoRI-BamHI digested plasmid pBR322 HuIGF-1 under standard ligation conditions mentioned above. Competent *E. coli* 294, prepared as above (3), were used as transformation hosts and the transformed cells were plated onto LB-Ampicillin agar plates. Several transformants were picked,

miniscreened as above (4), and two exhibiting an EcoRI-BamHI insertion were grown in large scale and their plasmids purified. Using the Maxam-Gilbert procedure (5) the construction was sequenced to verify the correct synthesis and insertion of the EcoRI-Sau96I collagenase linker. This plasmid was called pBR322 HuSynIGF-1-M.

To prepare this EcoRI-SalI IGF-1 coding sequence for insertion into pNCV and pNCVsLE, 30 μ g of pBR322 HuSynIGF-1-M was digested with 70 U of SalI in 200 μ l of 1X SalI buffer (150 mM NaCl, 6 mM Tris-HCl (pH 7.9); 6 mM $MgCl_2$; 6 mM 2-mercaptoethanol) at 37° for 6 hours. After addition of EDTA to 15 mM, the mixture was phenol extracted, chloroform extracted 2X, and ethanol precipitated.

Using standard chemical synthesis procedures (2) a SalI-EcoRI linker

5' TCGACGTACATG 3'
3' GCATGTACTTAA 5'

was synthesized and 400 pmols kinased, as above. 200 pmols of the kinased linker was ligated to the SalI digested pBR322 HuSynIGF-1-M (prepared above) with 800 U T4 DNA ligase in 30 μ l of 1X ligation buffer for 12-16 hours at 14°C.

After termination of the reaction with EDTA, the mixture was phenol extracted, chloroform extracted 2X, and ethanol precipitated. The pellet was then taken up in 1X EcoRI buffer and digested with 100 U EcoRI in a volume of 200 μ l for 8 hours at 37°. After addition of EDTA to a concentration of 15 mM, the mixture was chromatographed on a 6 percent polyacrylamide slab gel. The gel was stained and the ~230 bp band corresponding to the EcoRI-EcoRI HuIGF-1 fragment was extracted from the gel, phenol extracted, chloroform extracted 2X, and ethanol precipitated. This fragment was ready for ligation into pNCV and pNCVsLE. pNCV and pNCVsLE were prepared for ligation by digestion of 20 μ g of each with 100 U EcoRI in 200 μ l 1X EcoRI buffer at 37° for 8 hours. After digestion, 200 U of bacterial alkaline phosphatase was added to each reaction and the mixtures warmed to 65°C for 2 hours. EDTA

was added to a concentration of 15 mM and the mixes were phenol extracted 3X, chloroform extracted 2X and then ethanol precipitated. These expression vectors were prepared for ligation.

5 Ligations of the EcoRI-EcoRI Human IGF-1 fragment into the two expression vectors were performed in 30 μ l reaction volumes in 1X T4 DNA ligase buffer with 800 U T4 DNA ligase at 14° for 12-16 hours. The EcoRI-EcoRI fragment was present at a 10-fold molar excess to vector.

10 Competent E. coli 294 were prepared (3) (ATCC 31446) and used as transformation hosts for the ligations. Transformed cells were plated onto LB-agar plates containing tetracycline (5 μ g/ml; LB-Tet-plates) and transformants were miniscreened (4). Miniscreen plasmid DNA from transformants of the pNCV-IGF-1 construction were digested with both PstI
15 and BglII to determine the orientation of the EcoRI fragment insertions. Two clones whose plasmids contained a ~570 bp BglII-PstI fragment (as opposed to a ~690 bp fragment) were grown in large scale and their plasmids prepared. The construction was sequenced using the Maxam-Gilbert procedure (5) to confirm the correct insertion at the
20 junction of the trp fusion and IGF-1 protein coding sequences as well as retention of the desired reading frame. Plasmids with the correctly inserted IGF-1 fragment were called pNCVLE-IGF-1. Transformants of the pNCV-sLE-IGF-1 construction were also miniscreened by the same procedure (5), and the plasmid DNAs were digested with HincII and PstI. Two clones
25 exhibiting a ~150 bp HincII-PstI fragment (as opposed to a ~105 bp HincII-HincII fragment) were grown in large scale and their plasmids prepared. Using the Maxam-Gilbert techniques (5), the functions of the trp fusion and IGF-1 protein coding sequences were sequenced to ascertain proper orientation and retention of the proper reading frame. Those
30 plasmids possessing the correct insertion and proper reading frame were called pNCV-sLE-IGF-1.

To attempt expression of each of these constructions, two clones, one possessing pNCV-IGF-1 and the other possessing pNCV-sLE-IGF-1, were
35 inoculated into 10 ml M9-Tetracycline culture medium supplemented with

0.5 mg/ml Tryptophan. A clone containing pNCV-LE with no IGF-1 gene insert was also inoculated into culture medium to provide as a negative control in assays.

5 After 12-16 hours growth at 37° with agitation, 0.5 ml of these cultures were used to inoculate 250 milliliters of M9-Tetracycline culture medium. After growing for 12-16 hours at 37° with agitation, the cells were harvested by centrifugation at 5000 rpm for 10 minutes in a Sorvall GSA rotor. The refractile bodies were purified from the pelleted
10 cells by: a) suspending the host cells in a buffered solution of ionic strength suitable to solubilize most of the host protein, b) subjecting the suspension to cell wall/membrane disruption, c) centrifuging the disrupted suspension at low speed to form a pellet, optionally repeating the foregoing steps, and d) recovering the heterologous protein as
15 refractile bodies in the pellet (Reference 13). A small quantity of refractile particles of each of the three preparations was boiled in SDS and 2-mercaptoethanol containing sample buffer and run on SDS-polyacrylamide slab gels according to the Laemmli method (14). The size of the protein expressed by pNCV-IGF-1 (LE-IGF-1) was ~28,670
20 Daltons (see Figure 7), and ~9770 Daltons for the pNCV-sLE-IGF-1 protein (sLE-IGF-1) (see Figure 8). These two expressed proteins were subjected to solubilization in 6M Guanidine-HCl followed by 50-fold dilution with dilute buffers. The final buffer for pNCV-IGF-1 after dilution was 0.12 M Guanidine-HCl; .05 M Tris-HCl pH 8, 20 percent glycerol; 0.1 mg/ml BSA; .15 M NaCl; 0.1 mM EDTA and the final buffer after dilution of the
25 pNCV-sLE-IGF-1 refractile bodies was 0.14 M Guanidine-HCl; 25 mM Tris-HCl pH 7.6; 10 mM CaCl₂. After spinning out particulate matter, the two solutions containing solubilized trp-IGF-1 fusion proteins were assayed by a radioimmune assay procedure of Furlanetto et al. (23), as modified
30 by Hintz et al. (24). Both fusion proteins demonstrated activity in this assay. A negative control prep was also included in the assay and the control exhibited no measurable activity.

35

0390L/29

Expression and Secretion in Yeast

To avoid the necessity of refractile body purification and solubilization, from bacterial cell lysates, yeast expression-secretion systems were sought as an alternative. Aside from the advantage of avoiding protein purification from cell lysates, coupled expression-secretion systems might obviate a subsequent in vitro processing step to remove a fused protein. Available were three yeast expression-secretion systems. These were: 1) yeast α factor (22), employing yeast α -factor promoter and preprosequence; 2) yeast invertase (16) consisting of the invertase promoter and signal sequence; and 3) a hybrid, composed of the PGK promoter (25) and invertase signal (16).

Yeast Alpha-Factor Promoter Pre-Alpha Factor IGF-1 Plasmid Construction

To obtain expression of IGF-1 using the α factor promoter and preprosequence, a plasmid constructed by Singh (22) was used. Plasmid P65 (Fig. 9) possesses sequences of the α -factor promoter, α -factor preprosequence, yeast 2 micron terminator, the yeast Trp 1 gene, as well as portions of the pBR322 plasmid. Due to the dearth of convenient restriction sites in the α -factor preprosequence, to insert the IGF-1 coding sequence, the identical ~230 bp EcoRI-EcoRI HuSynIGF-1-M fragment that was ligated into pNCV and pNCVsLE (as mentioned previously in bacterial construction) was used. This EcoRI-EcoRI fragment contained the collagenase recognition site Proline-Alanine-Glycine-Proline, and allowed for collagenase digestion should IGF-1 be secreted as a fusion protein. The protein expressed in this construction (see Fig. 10) consists of the prepro α -factor protein fused to IGF-1.

To insert the ~230 bp EcoRI-EcoRI fragment, the plasmid P65 was partially digested in 1X EcoRI buffer with EcoRI, and then sized upon a 0.7 percent horizontal agarose gel. The band corresponding to the linearized singularly restricted plasmid was excised, eluted from the gel, and phenol extracted, chloroform extracted 2X, and then ethanol precipitated. This DNA pellet was then taken up in 50 mM Tris-HCl (pH 8)

and treated with bacterial alkaline phosphatase under conditions to ensure 100 percent dephosphorylation of the 5' protruding ends.

Following this treatment, the phosphatase activity was removed by first adding EDTA to a concentration of 15 mM, then extracting the DNA with phenol-3X, chloroform extracting 2X, and ethanol precipitating the vector. This material then contained linearized P65 vector, digested with EcoRI in either of two locations: one, either at the EcoRI site upstream of the α -factor promoter and preprosequence, or at another, at the EcoRI site just downstream of the α -factor promoter and preprosequence. The ~230 bp EcoRI-EcoRI IGF-1 fragment was ligated into the vector. The desired location of insertion was at the EcoRI site just downstream from the α -factor promoter and preprosequence.

The ligation was carried out under standard ligation conditions and the transformation hosts were competent E. coli 294 prepared according to Dagert and Ehrlich (3). The transformed cells were plated onto LB-Amp-Agar plates. Several transformants were miniscreened according to the method of Birnboim and Doly (4), and plasmid DNA prepared as such was digested both SalI and HindIII in the appropriate buffers. One of several clones which contained a plasmid with an ~110 bp EcoRI-HindIII fragment was grown in large scale and its plasmid was purified. This plasmid, YE9T α -factor EcoRI-EcoRI IGF-1 (see Fig. 11), was used to transform competent yeast strain 20B-12 (α trp pep4) cells according to the Hitzeman modification (19) of Hinnen et al. (17) and Beggs et al. (18) procedures.

Two such transformants, as well as a negative control transformant (with no IGF-1 insertion in the plasmid), were grown in suspension as were those of the yeast pre-invertase-IGF-1 plasmid transformations. Supernates were tested for secreted IGF-1 activity, as measured by the radioimmune assay procedure of Furlanetto et al. (23) as modified by Hintz et al. (24). Both supernates of transformants having plasmids with IGF-1 inserts contained IGF-1 activity and the negative control supernate did not. One of these transformants was grown in large scale in a 10 liter fermenter and the supernate contained secreted IGF-1 activity at a

peak level of $\sim 3\mu\text{g/ml}$. The IGF-1 activity of the fermentation supernate was also demonstrated by a placental membrane radioreceptor assay developed by Horner *et al.* (26).

5 Yeast Invertase Promoter Signal IGF-1 Plasmid Construction

Based upon evidence of correct processing and secretion in yeast of proteins with heterologous signal sequences (16), the yeast invertase expression-secretion system became of interest. Attempted first was expression of the yeast invertase signal protein fused to IGF-1 (Fig. 12), coupled with the processing and secretion of IGF-1, using the invertase promoter as a starting point for transcription.

The yeast invertase signal coding sequence was attached to the IGF-1 gene by the use of a NcoI-HindIII (~ 400 bp) fragment containing the initiation ATG codon and 5' end of the signal DNA sequence, and 4 DNA fragments synthesized by standard procedures (2):

```

5' AGCTTTCCTTTTCCTTTTGGC      3'
3'      AAGGAAAAGGAAAACCGACCAA 5'

5' TGGTTTTGCAGCCAAAATATCTGCAG 3'
3'      AACGTCGGTTTATAGACGTCCAG 5'

```

The construction began with the isolation of the 90 bp AvaII-BamHI IGF-1 left half fragment by AvaII digestion of a ~ 730 bp PvuI-BamHI fragment isolated from PvuI-BamHI digested pBR322-HuSynIGF-1.

After phosphorylation of all four synthetic DNA fragments using standard kination conditions, the four synthetic fragments were mixed with the AvaII-BamHI IGF-1 left half fragment and ligated using standard ligation conditions. Following inactivation of the ligase by phenol and chloroform extraction 2X, the ethanol precipitated DNA pellet was dissolved and digested with HindIII and BamHI in the appropriate buffers. Newly constructed HindIII-BamHI (ca. 140 bp) fragment was isolated and extracted from a 6 percent polyacrylamide gel. This material was then ligated into HindIII-BamHI digested pBR322 vector,

which had been first digested with HindIII, then BamHI in the appropriate buffers, followed by purification of the 4014 bp vector fragment from a 6 percent gel.

5 The transformation host was competent E. coli 294 prepared by standard procedures (3) and the transformed cells were plated onto LB-Ampicillin agar plates. Several transformants were miniscreened by the Birnboim-Doly procedure (4) and their plasmid DNAs digested with EcoRI and BamHI. Two plasmids containing a ~167 bp EcoRI-BamHI fragment
10 (illustrating the insertion of a 140 bp fragment into the HindIII and BamHI sites) were grown in large scale and their plasmids prepared. Using Maxam-Gilbert sequencing techniques (5), the entire 43 bp HindIII-AvaII section of DNA was sequenced to confirm the correct chemical synthesis and construction. The correctly constructed plasmid
15 was called pBR322-P-I-HuSynIGF HindIII-BamHI (~4154 bp).

To insert the right half of the IGF-1 gene, this newly created plasmid was digested with BamHI-SalI in the appropriate buffers and the larger fragment (~3879 bp) was purified by gel fractionation. pBR322
20 HuSynIGF was digested with BamHI-SalI in the appropriate buffers and the 115 bp BamHI-SalI fragment corresponding to the right half of the IGF-1 gene was isolated by gel fractionation. This 115 bp BamHI-SalI IGF-1 right half fragment was then ligated into the BamHI-SalI digested pBR322-P-I-IGF-1 LH HindIII-BamHI vector using standard ligation
25 conditions. Competent E. coli strain 294 prepared according to Dagert and Ehrlich (3) were used as transformation hosts and transformed cells were plated onto LB-Amp-Agar plates. Several transformants were miniscreened using standard techniques (4) and plasmid DNA prepared as such was digested with EcoRI and SalI in the appropriate buffers and
30 those plasmids illustrating an insertion of the BamHI-SalI fragment corresponding to the right half of IGF-1 were called pBR322 P-I-HuSynIGF-1 HindIII-SalI. One of the clones containing the pBR322 P-I-IGF-1 HindIII-SalI plasmid was grown in large scale and the plasmid was isolated. This plasmid was then digested with HindIII and SalI in
35 the appropriate buffer to prepare a 255 bp HindIII-SalI fragment

containing all of the IGF-1 gene and the 3' portion of the yeast invertase signal coding sequence. This fragment of DNA was isolated by polyacrylamide gel fractionation and prepared for ligation by standard techniques.

5

The (~400 bp) NcoI-HindIII fragment containing the 5' end of the DNA sequence coding for the invertase signal as well as the yeast invertase promoter was created by NcoI and HindIII digestion of plasmid YIpsp-LeIFA (16) in the appropriate buffers. The YIpsp-LeIFA plasmid was first
10 digested with NcoI to completion in the appropriate buffer, then phenol extracted, chloroform extracted 2X and ethanol precipitated. The linearized molecules were then taken up in 1X HindIII buffer and partially digested to generate the needed NcoI-HindIII (~400 bp) fragment which contains an internal HindIII restriction site. This NcoI-HindIII
15 fragment was then isolated by gel fractionation and prepared for ligation using standard techniques.

20

25

30

To provide for a vector, plasmid pUC12-YI (EcoRI-BamHI) (16) was digested with NcoI and Sall in the appropriate buffers. After
20 purification by gel fractionation, the ~2.6 kbp vector was eluted from the gel and prepared for ligation by standard techniques. To perform the final construction, a three-part ligation was arranged using standard ligation techniques. The DNA used in the ligation included the NcoI-Sall-digested pUC12-YI (EcoRI-BamHI) (16), the ~400 bp NcoI-HindIII and the ~255 bp HindIII-Sall fragments. After ligation, the material was transformed into competent E. coli 294 cells prepared according to Dagert
25 et al. (3). Transformed cells were plated onto LB-Amp-Agar plates and several transformants were miniscreened using the procedure of Birnboim and Doly (4). Plasmid DNA prepared as such was digested with NcoI and
30 Sall in the appropriate buffers and one of several clones containing plasmids exhibiting the insertion of a ~625 bp NcoI-Sall DNA fragment was grown in large scale and its plasmid was purified.

35

As a final step, this plasmid was linearized by digestion with Sall in the appropriate buffer. Sall-EcoRI linker, prepared as mentioned

above, and kinased under standard kination conditions, was ligated to the linearized vector to convert the Sall ends to EcoRI ends using standard ligation conditions. After termination of the ligation reaction by addition of EDTA to 15 mM, phenol extraction, chloroform extraction 2X and ethanol precipitation, the DNA pellet was dissolved in 1X EcoRI buffer, and digested with EcoRI. The EcoRI digestion released a ~1150 bp EcoRI fragment which contained the yeast invertase promoter, yeast invertase signal coding sequence and the IGF-1 coding sequence in one contiguous sequence. This material was isolated as a ~1150 bp band from a 6 percent polyacrylamide slab gel after fractionation and prepared for ligation using standard procedures.

The yeast-E. coli shuttle vector to receive this EcoRI fragment was prepared by EcoRI digestion of plasmid YEp9T (16) to linearize the vector, followed by treatment of the EcoRI termini with bacterial alkaline phosphatase using conditions recommended by the manufacturer to produce 100 percent dephosphorylation of the 5' protruding ends. The phosphatase reaction was terminated by addition of EDTA to 15 mM and the mixture phenol extracted 3X, chloroform extracted 2X, and then the DNA was ethanol precipitated. After redissolving the DNA pellet in 1X ligation buffer, the vector was mixed with the EcoRI ~1150 bp fragment and ligated under standard ligation conditions. Competent E. coli 294 cells prepared according to Dagert et al. (3) were used as transformation hosts and the transformants were plated onto LB-Amp-Agar plates. To determine the orientation of the insertion, several transformants were miniscreened using the method Birnboim and Doly (4) and plasmid DNAs purified as such were digested with BamHI in the appropriate buffer. One of several transformants possessing plasmids which produced a 1.3 kb BamHI-BamHI fragment upon BamHI digestion (as opposed to a ~475 bp fragment) was grown in large scale and its plasmid was purified. This plasmid, called P.I.IGF-1 EcoRI-EcoRI P.I. Promoter was used to transform competent yeast cells prepared essentially according to the methods of Hinnen, A., et al. (17), and Beggs, J.D. (18), but with the modifications of Hitzeman (19). The yeast strain 20B-12 (atrpl pep4) was used and was obtained from the Yeast Genetics Stock Center. In this construction, the

expression of IGF-1 begins with transcription at the invertase promoter and terminates in the yeast 2 micron sequence. The fusion protein expressed by this construction consisted of the yeast invertase signal fused to the IGF-1 protein, the combined molecular weight of which was 9964 Daltons. Another plasmid with the EcoRI fragment inserted in the reverse orientation was also used to transform competent yeast cells. In this construction, the IGF-1 was not provided with the yeast terminator.

Several transformants were picked and streaked on YNB-CAA agar plates. Among these, three transformants were picked and inoculated into 10 ml of YNB-CAA grow-up medium, in shake flasks. A fourth culture was also started using a colony transformed with the same vector, but with the EcoRI fragment inserted into the vector in the reverse orientation. After 16-20 hours growth at 30°, the cultures were sampled (1 ml) and cleared of cells by spinning 5' in an eppendorf microfuge. Supernatants were taken off and assayed for secreted activity using the radioimmune assay procedure of Furlanetto *et al.* (23) as modified by Hintz *et al.* (24). The supernates of the three transformants demonstrated activities of 1.7 to 3.3 ng/ml of IGF-1 activity and the negative control showed no activity. To determine intracellular activity, the pellets from 1 ml of culture were washed 1X in 25 mM Tris-HCl (pH 7.6), 1 mM EDTA and then lysed by 3-4 minutes of vigorous vortexing in 0.5 ml of the above Tris-EDTA solution with 0.4 ml of glass beads.

Assay of the cell lysates demonstrated IGF-1 activities of 1.5-2.8 ng/ml in the three IGF-1 secreting transformants and no activity in the negative control transformant. The highest secretor of the three transformants was grown in a 5 liter fermentation and the secreted IGF-1 activity reached a peak of 74 ng/ml of supernate.

Yeast PGK Promoter Pre-Invertase IGF-1 Plasmid Construction

One difficulty in the use of the invertase promoter was that it was subject to repression in the presence of glucose. Due to the incompatibility of glucose with high levels of transcription initiation

at the invertase promoter, the PGK promoter was sought as an alternative promoter, glucose, being the mainstay carbon source of fermentation processes.

5 To begin construction of the PGK promoter P.I.IGF-1 construction, it was necessary to clone a fragment containing the entire invertase signal coding sequence. To do this, plasmid pLeIF-A-Invertase Signal (16) was digested with BglII and then BamHI in the appropriate buffers. This digestion released several fragments, one of which was a ~625 bp
10 BglII-BamHI fragment which was isolated from a 6 percent polyacrylamide slab gel and prepared for ligation using standard techniques. To clone this fragment, the pUC8 vector (20) was chosen as a cloning vehicle. pUC8 plasmid was digested with BamHI in 1X BamHI buffer, treated with bacterial alkaline phosphatase to dephosphorylate the 5' termini, and
15 then run onto and purified from a 5 percent polyacrylamide slab gel.

After standard preparation for ligation the BamHI digested vector was mixed with the above ~625 bp BglII-BamHI fragment, and ligated under typical ligation conditions. The mixture was then transformed into
20 competent E. coli 294 prepared by the Dagert et al. method (3) and the transformed culture plated onto LB-Amp-Agar plates. Several transformants were picked and miniscreened using the Birnboim and Doly (4) technique. Miniscreen plasmid DNA was digested with EcoRI and an analytical gel of the digests illustrated two types of plasmids having EcoRI fragments
25 either ~260 bp or ~385 bp in length. One clone containing a ~260 bp EcoRI fragment was grown in large scale and its plasmid purified. This plasmid was called pUC8 P.I. Promotor-Signal BglII-BamHI.

A clone of this type was chosen because of the desired orientation of
30 the inserted BglII-BamHI fragment. What was needed from this plasmid was an ~20 bp EcoRI-HindIII fragment containing the ATG initiation codon and 5' end of the invertase signal coding sequence.

To construct the intact invertase signal coding DNA sequence, ~150 bp
35 HindIII-BamHI fragment containing the 3' end of the signal sequence fused

to the left half of the IGF-1 gene was isolated from HindIII-BamHI digestion of plasmid pBR322 P.I. IGF-LH HindIII-BamHI (~4154 bp). Isolation was by polyacrylamide slab gel fractionation, and the DNA band corresponding to the ~150 bp fragment was excised and prepared for
5 ligation using standard techniques.

To obtain the short (~20 bp) EcoRI-HindIII fragment, the plasmid pUC8 P.I. Promotor-Signal-BglII-BamHI was digested with EcoRI in 1X EcoRI buffer. This digestion released the ~260 bp EcoRI-EcoRI fragment which
10 was isolated from a 6 percent polyacrylamide slab gel after fractionation of the digestion mixture. This ~260 bp fragment was then digested with HindIII in the appropriate buffer, causing the creation of two HindIII-EcoRI fragments, one ~20 bp and the other ~240 bp in length. After complete digestion, the digestion was terminated by addition of
15 EDTA to 15 mM and the entire mix phenol extracted, chloroform extracted 2X, and then ethanol precipitated.

A vector was prepared by EcoRI-BamHI digestion of pBR322 (15) in the appropriate buffers followed by purification of the EcoRI-BamHI digested
20 vector from a 5 percent polyacrylamide slab gel. After preparation for ligation using standard techniques, the vector was mixed with the ~150 bp HindIII-BamHI fragment (3' end of invertase signal + Left Half IGF-1), and the two HindIII-EcoRI fragments (the ~20 bp fragment containing the 5' end of the invertase signal coding sequence), and the entire mixture
25 was ligated under standard ligation conditions. Competent E. coli 294 prepared according to Dagert and Ehrlich (3) were used as transformation hosts for the ligation, and the transformed cells plated onto LB-Amp-Agar plates. Several transformants were miniscreened according to Birnboim and Doly (4) and the purified miniscreen DNAs were digested with EcoRI
30 and BamHI. One of several clones possessing an ~170 bp EcoRI-BamHI fragment was grown in large volume and its plasmid purified. This plasmid contained the complete yeast invertase signal coding sequence fused to the left half of IGF-1 and was called P.I. IGF-1 L.H. RI-BamHI.

35 The desired ~170 bp EcoRI-BamHI fragment was isolated from this

plasmid by digestion of the plasmid with EcoRI and BamHI in the appropriate buffers followed by slab gel fractionation of the reaction mix. Using standard techniques, the ~170 bp band of DNA was prepared for ligation. To complete the construction, the right half of IGF-1 was isolated as an ~120 bp BamHI-EcoRI fragment from the plasmid P.I. IGF-1 EcoRI-EcoRI-P.I. Promoter by digestion with EcoRI and BamHI in the appropriate buffers followed by elution from a gel slice after polyacrylamide slab gel fractionation of the digestion mixture. These two fragments, the ~170 bp EcoRI-BamHI and the ~120 bp BamHI-EcoRI, were ligated together in vitro under standard ligation conditions, with both fragments present in roughly equimolar concentrations. This ligation mixture was then terminated by the addition of EDTA to ~15 mM followed by phenol extraction, chloroform extraction 2X, and ethanol precipitation. The DNA pellet was then taken up in 1X EcoRI buffer and digested with EcoRI. The digest was then run on a 6 percent polyacrylamide slab gel and the DNA band staining at ~290 bp (as opposed to ~340 bp and 240 bp) was excised and prepared for ligation using standard techniques. This ~290 bp EcoRI-EcoRI fragment contained the entire yeast invertase signal coding sequence fused to the complete IGF-1 coding sequence.

To express this protein, it was necessary to select a yeast vector with a promoter. The PGK promoter of the plasmid YEplPT Small (see Fig. 13) was used. YEplPT Small was constructed as a derivative of YEplPT (21) by ClaI and PvuII digestion of YEplPT in the appropriate buffers. The ClaI 5' protruding end was converted to a blunt end by use of DNA polymerase I (Klenow) under conditions recommended by the vendor. After blunting the ClaI protruding ends, the blunt ends ClaI and PvuII of the linearized vector were fused using T4 DNA ligase under standard ligation conditions. The resultant YEplPT small vector was ~5.9 kbp in size (or ~2.7 kbp smaller than YEplPT). Just as YEplPT, YEplPT small possesses the 2 micron origin and terminator, the PGK promoter, the TRP1 gene, and sequences from pBR322, including the β -lactamase gene.

YEplPT Small was employed as a vector by insertion of the ~290 bp EcoRI fragment into the unique EcoRI site of the plasmid. EcoRI

linearized YEplPT Small vector was prepared by EcoRI digestion of YEplPT small followed by bacterial alkaline phosphatase (BAP) treatment (to prevent religation of the complementary termini). The BAP was removed by phenol extraction 3X, chloroform extraction 2X, and ethanol precipitation. Under standard ligation conditions, the ~290 bp EcoRI fragment was ligated into the vector.

Competent *E. coli* 294 prepared according to Dagert and Ehrlich (3) were used as transformation hosts and the transformed culture was plated onto LB-Amp-Agar plates. Several transformants were miniscreened by the Birnboim and Doly procedure (4) and miniscreen plasmid DNAs were digested with HindIII in the appropriate buffer to determine the orientation of the insert. One of several transformants possessing a plasmid with a ~400 bp HindIII fragment was grown in large scale and its plasmid was purified. This plasmid was called YEplPT Small P.I. IGF-1 PGK promoter (see Fig. 14) and was used to transform competent yeast strain 20B-12 (ATCC 20626) (α trp pep4) cells employing the Hitzeman modification (19) of Hinnen *et al.* (17), and Beggs *et al.* (18) procedures.

Several yeast transformants were grown in suspension in identical fashion as were those of the P.I. IGF-1 EcoRI-EcoRI P.I. promoter plasmid transformation and supernates were measured for activity determined by a radioimmune assay method of Furlanetto *et al.* (23) as modified by Hintz *et al.* (24). Shake flask supernates of three transformants contained activities ranging from 38 to 53 ng/ml of supernate. Similarly, one of these transformants was selected and grown in larger scale, utilizing a 10 liter fermenter and the secreted IGF-1 activity in the supernate reached a peak of ~780 ng/ml. This fermentation supernate was also subjected to a radioreceptor assay (26) and was demonstrated to contain IGF-1 activity.

Mature Human IGF Production

To construct a DNA sequence coding for the α -factor pre-pro protein fused to the DNA sequence coding for mature IGF-I, an M-13 *in vitro*

mutagenesis technique was employed. (See Regin *et al.*, *Proc. Acad. Science (USA)* **75**, 4268; Hutchinson, *et al.*, *Journal Biological Chem.* **253**, 5551; Gilliam, *et al.*, *Gene* **8**, 81 and 99; Gillam, *et al.*, *Nucleic Acids Research* **6**, 2973; Adelman, *et al.*, *DNA* (June, 1983).)

5

To construct the M-13 plasmid, the plasmid YEp9T α -factor ECoRI-ECoRI IGF-I (Figure 16) was digested with Bgl II and Sal I and the ca. 1.5 Kbp fragment containing the α -factor promoter-signal fused to IGF-I was isolated by polyacrylamide gel electrophoresis. This fragment was then ligated under standard ligation conditions to an MP-8 (BRL) vector digested with BamHI and Sal I, and treated with bacterial alkaline phosphatase. This ligation mix was then transformed into competent JM101 cells prepared according to the method of Dagert and Ehrlich (3). These transformants were then mixed with non-competent JM101 cells grown to log phase, mixed with top agar and plated onto LB agar plates. Several clear plaques were picked and sequenced using the M-13 dideoxy sequencing technique to confirm the presence of an insertion into the SalI-BamHI sites of the vector.

20

To perform the deletion according to the method above, a single strand of DNA of the sequence

5' AGAGTTTCCGGACCT CTT TTATCCAAAG 3'

25

was chemically synthesized by standard methods (2) and used to delete the DNA sequence

5' GAGGCTGAAGCTCTAGAATTCCTGCC 3'

3' CTCCGACTTCGAGATCTTAAGGGACGG 5'

30

just preceding the IGF-1 coding sequence of the α -factor promoter/signal IGF-I fusion sequence. This construction was then isolated as a replicative form, using a large scale plasmid preparation procedure from a JM101 cell culture inoculated with this plasmid containing the deletion.

35

The isolated replicative form (10 mg) was then digested with SalI.

Then phenol-chloroform extracted and then ethanol precipitated and prepared for ligation. To this replicative form was ligated Sal I-EcoRI linkers. After ligation and inactivation of the ligase by phenol, chloroform extraction followed by ethanol precipitation, the material was
5 digested with ~50 U EcoRI enzyme under standard conditions and then run onto a 6 percent polyacrylamide gel. The ca. 1.5 kbp RI-EcoRI fragment released was isolated from the gel and prepared for ligation using standard conditions.

10 Yeast vector was prepared by digestion of 10 mg YEP9T plasmid with 50 units of EcoRI followed by treatment with bacterial alkaline phosphatase. The digestion was then repeatedly phenol-chloroform extracted and then ethanol precipitated and prepared for ligation.

15 The ca. 1.5 kbp EcoRI-EcoRI fragment containing the deletion was then ligated to the EcoRI-EcoRI YEP9T vector and the ligation mix was then transferred into competent 294 cells prepared according to the method of Dagert and Erhlich (3) and miniscreened using the method of Birnboim and Doly (4). DNA prepared was screened by digestion with EcoRI and those
20 DNAs illustrating an insertion of the ca. 1.5 kbp fragment were used to transform competent yeast strain 203-12 (ATCC 20526) according to the modification of Hitzerman (19) of the Hinner, et al., (17), and Beggs, et al., (18) procedures.

25 Transformants were then grown in shaker flasks and supernates assayed and shown to have IGF-I activity by the radioimmune assay procedure of Furlanetto, et al., (23) as modified by Hintz, et al., (24).

30 One of these clones were grown in large scale in a 10-liter fermentor and IGF-I purified from the supernatant of this fermentation. This material was then subjected to amino terminal protein sequencing and shown to be mature IGF-I protein.

35 Human EGF is prepared in accordance with invention following analogous procedures as those described above.

Construction, Expression, and Secretion of Human EGF

In a fashion similar to IGF-1, double stranded DNA (Fig. 15) synthesized either by chemical means or through polymerization reactions was assembled to form a mature EGF coding sequence, with a codon coding for methionine (ATG) just preceding the amino-terminal asparagine found in the mature protein, and a codon (GTC) substituting valine for methionine at residue number 21 from the amino-terminal asparagine. This construction was then attached at the 5' end to an additional coding sequence, which when expressed in yeast or bacteria produced a fusion protein. This fusion protein was then susceptible to CNBr cleavage at the methionine to release the valine substituted human EGF molecule.

To secrete the mature form of EGF from yeast, the above sequence coding for the mature protein was attached to the α -factor promoter/prepro sequence, the codon coding for valine at residue number 21 was replaced by ATG, and the appropriate deletion was made to bring the coding sequence for mature EGF adjacent to the α -factor signal coding sequence (Figure 16). This construction was then inserted into the yeast vector Yep9T and transformed into yeast. Transformants produced as such expressed and secreted mature human EGF. In addition, the sequence coding for mature EGF was attached to the preinvertase signal sequence (Figure 17) and this construction, when inserted into the yeast vector Yep1PT small containing the PGK promoter, and transformed into yeast, resulted in the expression and secretion of human EGF.

Construction, Expression, and Secretion of Human IGF-II

A double stranded DNA sequence coding for mature IGF-II was constructed from a combination of synthetic and natural DNA sequences (Figure 18). This coding sequence, which did not contain an internal methionine, was attached to the TrpE leader protein coding sequence and was expressed as a fusion protein. Mature IGF-II was chemically cleaved from the purified fusion product by the action of CNBr upon a methionine residue preceding the first residue (alanine) of the mature protein.

The IGF-II coding sequence was also attached to the α -factor promoter/prepro sequence and after the appropriate deletion was made to bring the 3' end of the α -factor signal coding sequence adjacent to the 5' end of mature IGF-II coding sequence, the construction was inserted into the Yep9T vector and transformed into yeast. Resultant transformants expressed and secreted mature human IGF-II. In the same manner, the sequence coding for mature IGF-II was attached to the preinvertase coding sequence. The resultant construction was inserted into YepIPT small and transformed into yeast. Transformants produced as such expressed and secreted mature human IGF-II.

Pharmaceutical Compositions

The compounds of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the human IGF and human EGF or products hereof are combined in admixture with a pharmaceutically acceptable carrier vehicle. Suitable vehicles and their formulation, inclusive of other human proteins, e.g. human serum albumin are described, for example, in Remington's Pharmaceutical Sciences by E. W. Martin, which is hereby incorporated by reference. Such compositions will contain an effective amount of the protein hereof together with a suitable amount of vehicle in order to prepare pharmaceutically acceptable compositions suitable for effective administration.

Notwithstanding that reference has been made to particular preferred embodiments of the present invention, it will be understood that the present invention is not to be construed as limited to such rather to the lawful scope of the appended claims.

Bibliography

- A. Rinderknecht, E. W. et al., Proceedings National Academy of Sciences (USA) 73, 2365 (1976).
- 5 B. Rinderknecht, et al., Proceedings National Academy of Sciences (USA) 73, 4379 (1976).
- C. Blundell, et al., Proceedings National Academy of Sciences (USA) 75, 1980 (1978).
- 10 D. Rinderknecht, et al., Journal of Biological Chemistry 253, 2769 (1978), 2365 (1976).
- E. Rinderknecht, et al., FEBS Letters 89, 283 (1978).
- F. Zaph, et al., Metabolism 27, 1803 (1978).
- G. Hintz, et al., Journal of Clinical Endocrinology and Metabolism 50, 405 (1980).
- 15 H. Blundell, et al., Nature 287, 781 (1980).
- I. Hintz, et al., Journal of Clinical Endo. and Metabolism 51, 672 (1980).
- J. Baxter, et al., Journal of Clinical Endo. and Metabolism 54, 474 (1980).
- 20 K. Hintz, et al., Journal of Clinical Endo. and Metabolism 54, 442 (1982).
- L. Schoenle, et al., Nature 296, 252 (1982).
- M. British Patent Application Publication No. 2007676A.
- N. Wetzel, American Scientist 68, 664 (1980).
- 25 O. Microbiology, Second Edition, Harper and Row Publications Inc., Hagerstown, Maryland (1973), especially pages 1122 et sequence.
- P. Scientific American 245, 106 (1981).
- 30 1. Rinderknecht, E. and Humbel, R.E., J. Biol. Chem. 253, 8, 2769-2776 (1978).
2. Crea, R. and Horn, T., Nucleic Acids Research 8, 2331-2348 (1980).
3. Dagert, M. and Ehrlich, S.D., Gene 6, 23-28 (1979).
4. Birnboim, H.C. and Doly, J., Nucleic Acids Research 7, 1513-1523 (1979).
- 35 5. Maxam, A. and Gilbert, W., Methods in Enzymology 65, 499 (1980).

6. Villa-Komaroff et al., Proc. Natl. Acad. Sci. USA 75, 3727 (1979).
7. Rowenkamp and Firtel, Dictyostelium Dev. Biol. 79, 409 (1980).
8. Munsch, E. et al., Hoppe-Seyler's Z. Physiol. Chem. Bd. 362,
S1285-1287 (Sept. 1981).
- 5 9. Maniatis, T. et al., Molecular Cloning, 426 (1982).
10. Kleid, D.G., Annals of New York Acad.Sci., Biochem.Eng., II, 413, 23-30,
(1983).
11. Seifter, S. and Gallop, P.M., The Proteins, 2nd Ed. (H. Neurath, ed.)
Vol. V, p. 659 (1966).
12. Nordwig, A., Leder 13, 10 (1962).
- 10 13. Lin, N. (U.S.S.N. 06/452,363, filed 22.12.82.) EPA 83307840.5.
14. Laemmli, U.K. Nature (London) 227, 680-685 (1970).
15. Bolivar, F. et al., Gene 2, 95 (1977).
16. Chang, C.N., (U.S.S.N. 06/488337 filed 25.4.83) EPA 84302722.8.
17. Hinnen, A. et al., Proc. Natl. Acad. Sci. USA 75, 1929-1933 (1978).
- 15 18. Beggs, J.D., Nature 275, 104-109 (1978).
19. EP 88632.
20. EP 88622.
- 20 21. Hitzeman, R.A. et al., Science 219, 620-625 (1983).
22. Singh, A. (U.S.S.N. 06/488323 filed 25.4.83) EPA 84302723.6.
23. Furlanetto, R.W., Underwood, L.E., Van Wyk, J.J., D'Ercole, A.J., J.
Clin. Invest. 60, 648 (1977).
24. Hintz, R.L., Liu, F., Marshall, L.B., Chung, D., J. Clin. Endocrinol.
Metab. 50, 405 (1980).
- 25 25. Hitzeman, R.A. et al., Proceedings of the Berkeley Workshop on Recent
Advances in Yeast Molecular Biology: Recombinant DNA. U.C. Press,
Berkeley, p. 173 (1982).
26. Horner, J.M., Liu, F., Hintz, R.A., J. Clin. Endocrinol. Metab. 47,
6, p. 1287 (1978).
- 30 27. Rossi, J.J., Zierzek, R., Huang, T., Walker, P.A., Itakura, K., J.
Biol. Chem. 257, 16, 9226 (1982).

CLAIMS

1. Mature human IGF, as a product of expression, processing, and secretion from a recombinant host.
- 5 2. Mature human IGF, as a product of recombinant yeast host expression with a N-terminus pre-sequence derived from yeast alpha factor DNA, processing of said pre-sequence and secretion of the mature human IGF into the medium supporting the recombinant yeast host.
- 10 3. Human IGF produced by recombinant host.
4. Human EGF produced by recombinant host.
- 15 5. A process for producing human IGF comprising preparing a replicable expression vector capable of expressing the DNA sequence encoding human IGF in a suitable host cell, transforming a host cell culture with said vector to obtain a recombinant host cell, culturing said recombinant host cell culture under conditions permitting expression
20 of said human IGF encoded DNA sequence to produce human IGF and recovering said human IGF.
- 25 6. A process for producing human EGF comprising preparing a replicable expression vector capable of expressing the DNA sequence encoding human EGF in a suitable host cell, transforming a host cell culture with said vector to obtain a recombinant host cell, culturing said recombinant host cell culture under conditions permitting expression
30 of said human EGF encoded DNA sequence to produce human EGF and recovering said human EGF.
- 35 7. An expression vector capable, in a transformant recombinant host, of expressing a DNA sequence encoding human IGF.

8. An expression vector capable, in a transformant recombinant host, of expressing a DNA sequence encoding human EGF.
9. A recombinant host transformed with the vectors of Claims 8 or 9.
10. A pharmaceutical composition consisting essentially of mature human IGF.

Chemically Synthesized DNA - IGF-1

Fig.1.

IGF-1 LEFT HALF

1L (43-mer)

AGTTCTGATTTCGAATTCATGGGTCCGGAAACTCTGTGCGGC
GACACGCCGCGACTCGACCAACTGCGAGACGTCTTAGTCTTGA

3L (43-mer)

2L (46-mer)

AGTTCTGATTCTGCAGTTCGTATGTGGTATCGAGGCTTCTACTTC
GAAGATGAAGTTGTTTGGCTGACCCATGCCTAGGAGTTAGTCTTGA

4L (46-mer)

IGF-1 RIGHT HALF

1R (46-mer)

GACTGACTTCTGGATCCTCTCTCGTGTGCTCCGCAAAACCGGCAT
TTGGCCGTAGCAACTACTTACGACAAAAGCCAGGACATTTCAGTCAG

3R (46-mer)

2R (54-mer)

TGACTGACTTGGTCTGTGACCTTCGCCGTCTGGAAATGTACTGCGCTCCGCTG
CGAGGCGACTTTGGCCGATTTCAGACGTATCAGCTGAGTTCAGTCAG

4R (46-mer)

DNA Polymerase I Synthesized Double Stranded DNAs

Fig.2.

IGF-1 LEFT HALF

12 additional bases
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
IniGlyProGluThrLeuCysGlyAlaGluLeuValAspAlaLeuGln
AGTTCTGATTTCGAATTCATGGGTCCGGAAACTCTGTGCGGCGCTGAGCTGGTTGACGCTCTGCAGAAATCAGAACT
TCAAGACTAAAGCTTAAGATACCCAGGCCCTTTGAGACACGCCGCGGACTCGACCAACTGCGAGACGCTCTTAGTCTTGA
PstI
EcoRI

Part 1

11 additional bases
15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
GlnPheValCysGlyAspArgGlyPheTyrPheAsnLysProThrGlyTyrGly
AGTTCTGATTCTGGCAGTTCTGATGTTGATCGAGGCTTCTACTTCAACAACCGACTGGGTACGGATCCTCAATCAGAACT
TCAAGACTAAGACGCTCAAGCATACACCACCTAGCTCCGAGATGAAGTTGTTTGGCTGACCCATGCTAGGAGTCTTGA
BstI
BamHI

Part 2

IGF-1 RIGHT HALF

11 additional bases
33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51
SerSerSerArgArgAlaProGlnThrGlyIleValAspGluCysCysPheArgSer
GACTGACTTCTGGATCCTCTCGTCTGCTCCGAAACCGGATCGTTGATGATGCTGTTTCGGTCCCTGTAAAGTCAGTC
CTGACTGAAGACCTAGGAGGAGAGCAGCAGGCGGTTTGCCCGTAGCACTACTTACGACAAAAGCCAGGACATTTCAGTCAG
BamHI
AvaII

Part 3

10 additional bases
51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70
SerCysAspLeuArgArgLeuGluMetTyrCysAlaProLeuLysProAlaLysSerAlaTer
TGACTGACTTGGTCTCTGACCTTCGCCGCTCGGAAATGATCTGCGCTCCGCTGAACCGGCTAAGCTCTGCATAGTCGACTCAAGTCAGTC
ACTGACTGAACCCAGGACACTGGAAGCGGCAGACCTTTACATGACCGGAGCGGACCTTTGGCCGATTTCAGACGTATCAGCTGAGTTTCAGTCAG
AvaII
SalI

Part 4

IGF-1 LEFT HALF

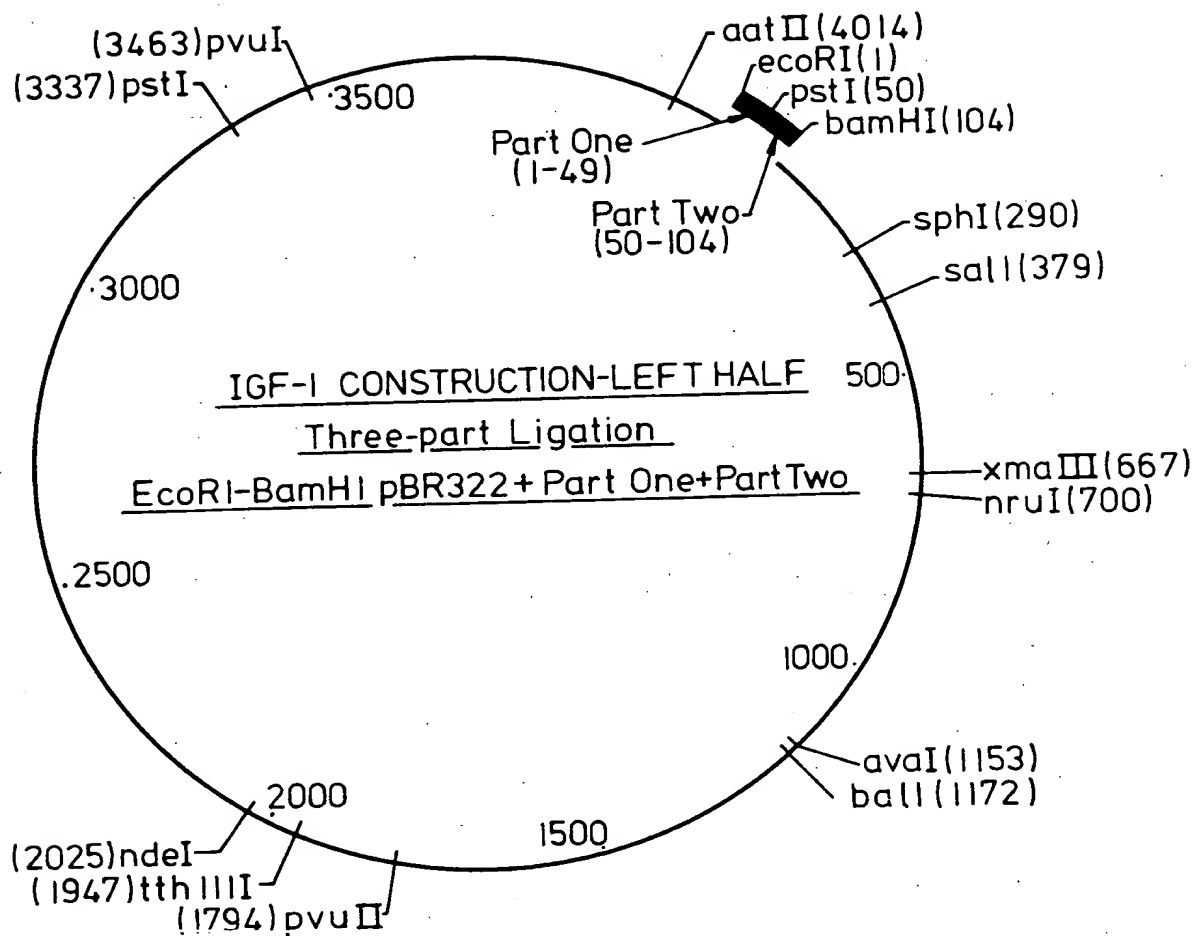
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
 IniGlyProGluThrLeuCysGlyAlaGluLeuValAspAlaLeuGln
AATTCTATGGTCCGGAACCTCTGTCCGGCGCTGAGCTGGTTGACGCTCTGCAG
GATACCCAGGCCTTTGAGACACGCCGCGACTCGACCAACTGCGAGA
ECORI PstI

Part 1

16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
 PheValCysGlyAspArgGlyPheTyrPheAsnLysProThrGlyTyr
TTCGTATGTGGTGATCGAGGCTTCTACTTCAACAAACCGACTGGGTAC
CGTCAAGCATACACCACCTAGCTCCGAAGATGAAGTTGTTGGCTGACCCCATGCCTAG
PstI BamHI

Part 2

Fig. 3.

*Fig.4.*

IGF-1 RIGHT HALF

33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51
 SerSerSerArgArgAlaProGlnThrGlyIleValAspGluCysCysPheArgSer
GACTGACTTCTGGATCCTCTCTCGTCGTGCTCCGCAACCCGGCATCGTTGATGAATGCTGTTTCGGTCTGTAAGTCAGTCGCCCCCCCCCCCCCCC
 Part 3 CCCCCCCCCCCCCCTGACTGAAGACCCTAGGAGGAGAGCAGCACCAGGGCGTTTGGCCGTAGCAACTACTTACGACAAAAGCCAGGACATTTCAGTCAG
BamHI AvaII

51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70
 SerCysAspLeuArgArgLeuGluMetTyrCysAlaProLeuLysProAlaLysSerAlaTer
TGACTGACTTGGTCTTGACCTTCGCCGTCTGGAAATGTACTGGCTCCGCTGAACCCGGCTAAGTCTGCATAGTCGACTCAAGTCAGTCGCCCCCCCCCCCCCCC
 Part 4 CCCCCCCCCCCCCCTGACTGAACCCAGGACACTGGAAGCGGCAGACCTTTACATGACGCGAGGCGACTTTGGCCGATTTCAGACGTATCAGCTGAGTTCAGTCAG
AvaII SaII

Fig. 5.

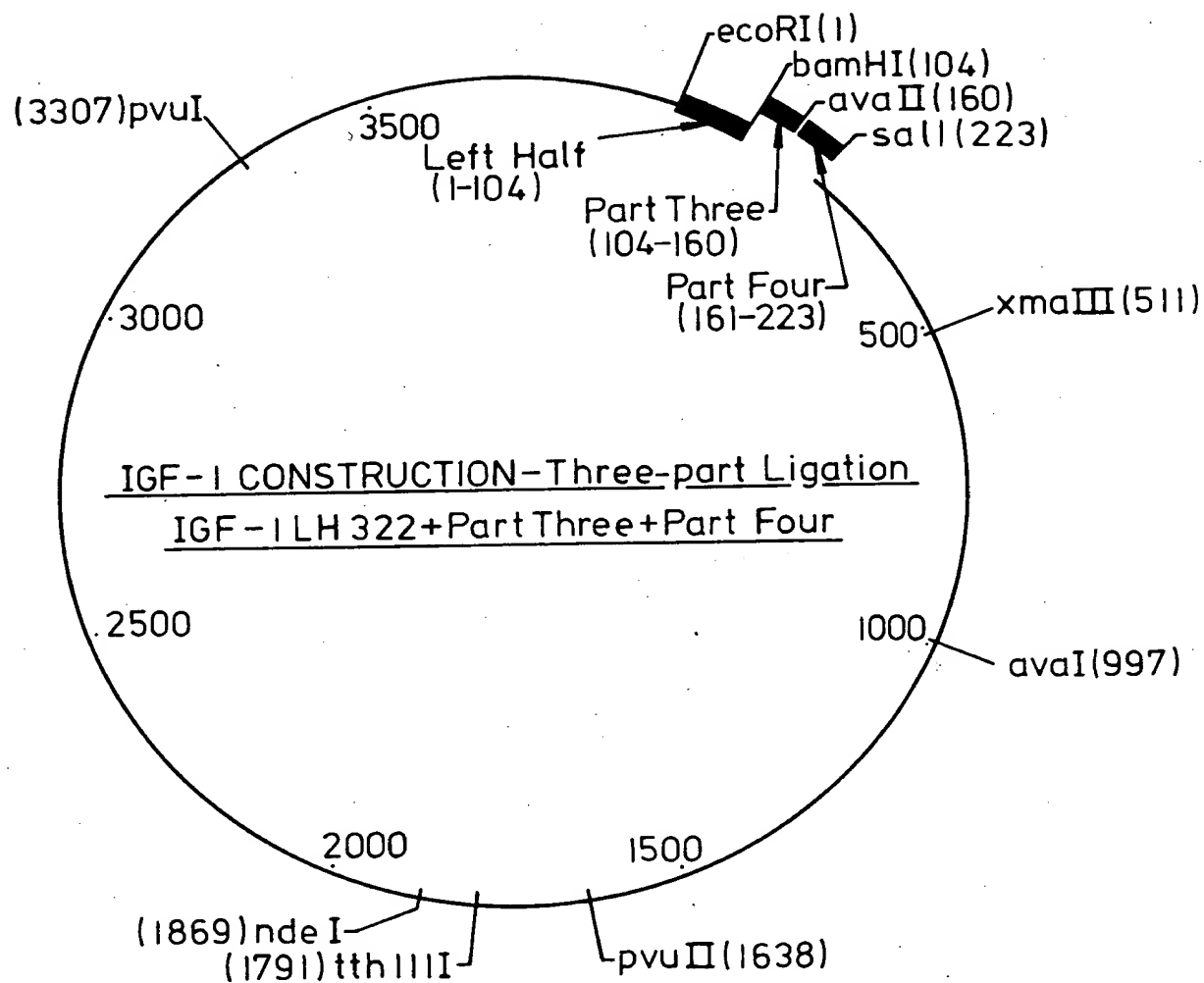
*Fig. 6.*

Fig. 7.

met lys ala ile phe val leu lys gly ser leu asp arg asp leu asp ser arg ile glu
 ATG AAA GCA ATT TTC GTA CTG AAA GGT TCA CTG GAC AGA GAT CTC GAC AGC CGT ATT GAA
 leu glu met arg thr asp his lys glu leu ser glu his leu met leu val asp leu ala
 CTG GAA ATG CGT ACC GAT CAT AAA GAG CTG TCT GAA CAT CTG ATG CTG GTT GAT CTC GCC
 arg asn asp leu ala arg ile cys thr pro gly ser arg tyr val ala asp leu thr lys
 CGT AAT GAT CTG GCA CGC ATT TGC ACC CCC GGC AGC CGC TAC GTC GCC GAT CTC ACC AAA
 val asp arg tyr ser tyr val met his leu val ser arg val val gly glu leu arg his
 GTT GAC CGT TAT TCC TAT GTG ATG CAC CTC GTC TCT CGC GTA GTC GGC GAA CTG CGT CAC
 asp leu asp ala leu his ala tyr arg ala cys met asn met gly thr leu ser gly ala
 GAT CTT GAC GCC CTG CAC GCT TAT CGC GCC TGT ATG AAT ATG GGG ACG TTA AGC GGT GCG
 pro lys val arg ala met gln leu ile ala glu ala glu gly arg arg arg gly ser tyr
 CCG AAA GTA CGC GCT ATG CAG TTA ATT GCC GAG GCG GAA GGT CGT CGC CGC GGC AGC TAC
 gly gly ala val gly tyr phe thr ala his gly asp leu asp thr cys ile val ile arg
 GGC GGC GCG GTA GGT TAT TTC ACC GCG CAT GGC GAT CTC GAC ACC TGC ATT GTG ATC CGC
 ser ala leu val glu asn gly ile ala thr val gln ala gly ala gly val val leu asp
 TCG GCG CTG GTG GAA AAC GGT ATC GCC ACC GTG CAA GCG GGT GCT GGT GTA GTC CTT GAT
 ser val pro gln ser glu ala asp glu thr arg asn lys ala arg ala val leu arg ala
 TCT GTT CCG CAG TCG GAA GCC GAC GAA ACC CGT AAC AAA GCC CGC GCT GTA CTG CGC GCT
 ile ala thr ala his his ala gln glu phe pro ala gly pro glu thr leu cys gly ala
 ATT GCC ACC GCG CAT CAT GCA CAG GAA TTC CCT GCC GGT CCG GAA ACT CTG TGC GGC GCT
 glu leu val asp ala leu gln phe val cys gly asp arg gly phe tyr phe asn lys pro
 GAG CTG GTT GAC GCT CTG CAG TTC GTA TGT GGT GAT CGA GGC TTC TAC TTC AAC AAA CCG
 thr gly tyr gly ser ser ser arg arg ala pro gln thr gly ile val asp glu cys cys
 ACT GGG TAC GGA TCC TCC TCT CGT CGT GCT CCG CAA ACC GGC ATC GTT GAT GAA TGC TGT
 phe arg ser cys asp leu arg arg leu glu met tyr cys ala pro leu lys pro ala lys
 TTT CCG TCC TGT GAC CTT CGC CGT CTG GAA ATG TAC TGC GCT CCG CTG AAA CCG GCT AAG
 ser ala AM
 TCT GCA TAG TCGACGTACATGAA

Translated Mol. Weight = 28671.51

Underscored - Mature Human IGF-1

Boxed - Collagenase Recognition Site

Fig. 8.

1 10
met lys ala ile phe val leu lys gly ser leu asp arg asp pro glu phe
ATG AAA GCA ATT TTC GTA CTG AAA GGT TCA CTG GAC AGA GAT CCA GAA TTC

20 30
pro ala gly pro glu thr leu cys gly ala glu leu val asp ala leu gln phe val cys
CCT GCC GGT CCG GAA ACT CTG TGC GGC GCT GAG CTG GTT GAC GCT CTG CAG TTC GTA TGT

40 50
gly asp arg gly phe tyr phe asn lys pro thr gly tyr gly ser ser ser arg arg ala
GGT GAT CGA GGC TTC TAC TTC AAC AAA CCG ACT GGG TAC GGA TCC TCC TCT CGT CGT GCT

60 70
pro gln thr gly ile val asp glu cys cys phe arg ser cys asp leu arg arg leu glu
CCG CAA ACC GGC ATC GTT GAT GAA TGC TGT TTT CGG TCC TGT GAC CTT CGC CGT CTG GAA

80 89
met tyr cys ala pro leu lys pro ala lys ser ala AM
ATG TAC TGC GCT CCG CTG AAA CCG GCT AAG TCT GCA TAG

Boxed – Collagenase Recognition Site

9/18

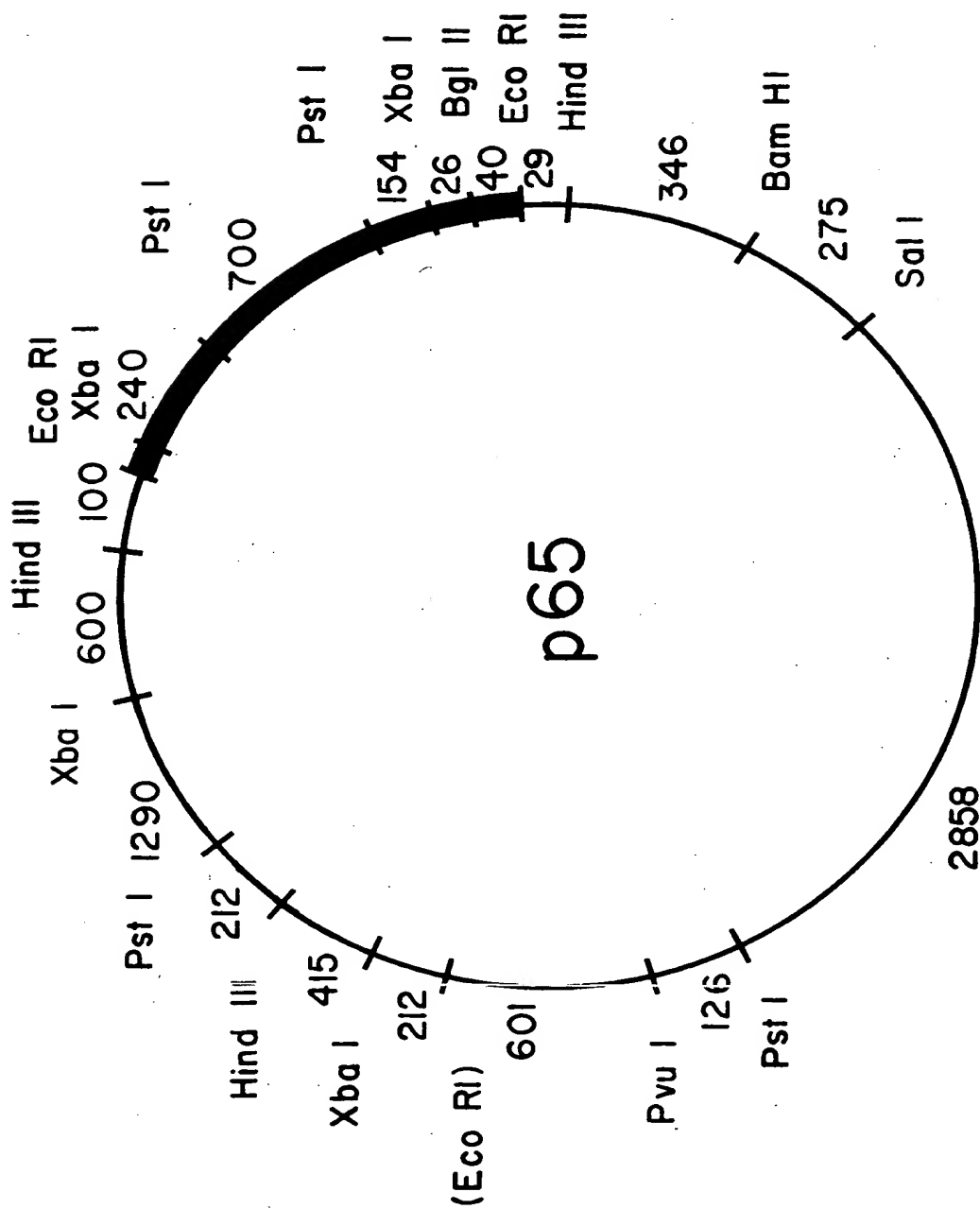


Fig. 9.

10/18

Yeast prepro α factor - IGF-1*Fig. 10.*

met arg phe pro ser ile phe ser ala val leu phe ala ala ser ser ala leu ala ala
 ATG AGA TTT CCT TCA ATT TTT AGT GCA GTT TTA TTC GCA GCA TCC TCC GCA TTA GCT GCT

pro val asn thr thr thr glu asp glu thr ala gln ile pro ala glu ala val ile gly
 CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA CAA ATT CCG GCT GAA GCT GTC ATC GGT

tyr ser asp leu glu gly asp phe asp val ala val leu pro phe ser asn ser thr asn
 TAC TCA GAT TTA GAA GGG GAT TTC GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT

asn gly leu leu phe ile asn thr thr ile ala ser ile ala ala lys glu glu gly val
 AAC GGG TTA TTG TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA

ser leu asp lys arg glu ala glu ala leu glu phe pro ala gly pro glu thr leu cys
 TCT TTG GAT AAA AGA GAG GCT GAA GCT CTA GAA TTC CCT GCC GGT CCG GAA ACT CTG TGC

gly ala glu leu val asp ala leu glu phe val cys gly asp arg gly phe tyr phe asn
 GGT GCT GAA CTG GTT GAC GCT CTG GAG TTC GTA TGT GGT GAC CGT GGC TTT TAC TTC AAC

lys pro thr gly tyr gly ser ser ser arg arg ala pro gln thr gly ile val asp glu
 AAA CCG ACT GGT TAC GGA TCC TCC TCT CGT CGC GCT CCG CAA ACT GGC ATC GTT GAT GAA

cys cys phe arg ser cys asp leu arg arg leu glu met tyr cys ala pro leu lys pro
 TGC TGT TTT CGT TCT TGT GAC CTG CGC CGT CTG GAA ATG TAC TGC GCT CCG CTG AAA CCG

ala lys ser ala Term
 GCT AAG TCT GCA TAG

Underscored - Mature Human IGF-1

Boxed - Collagenase Recognition Site

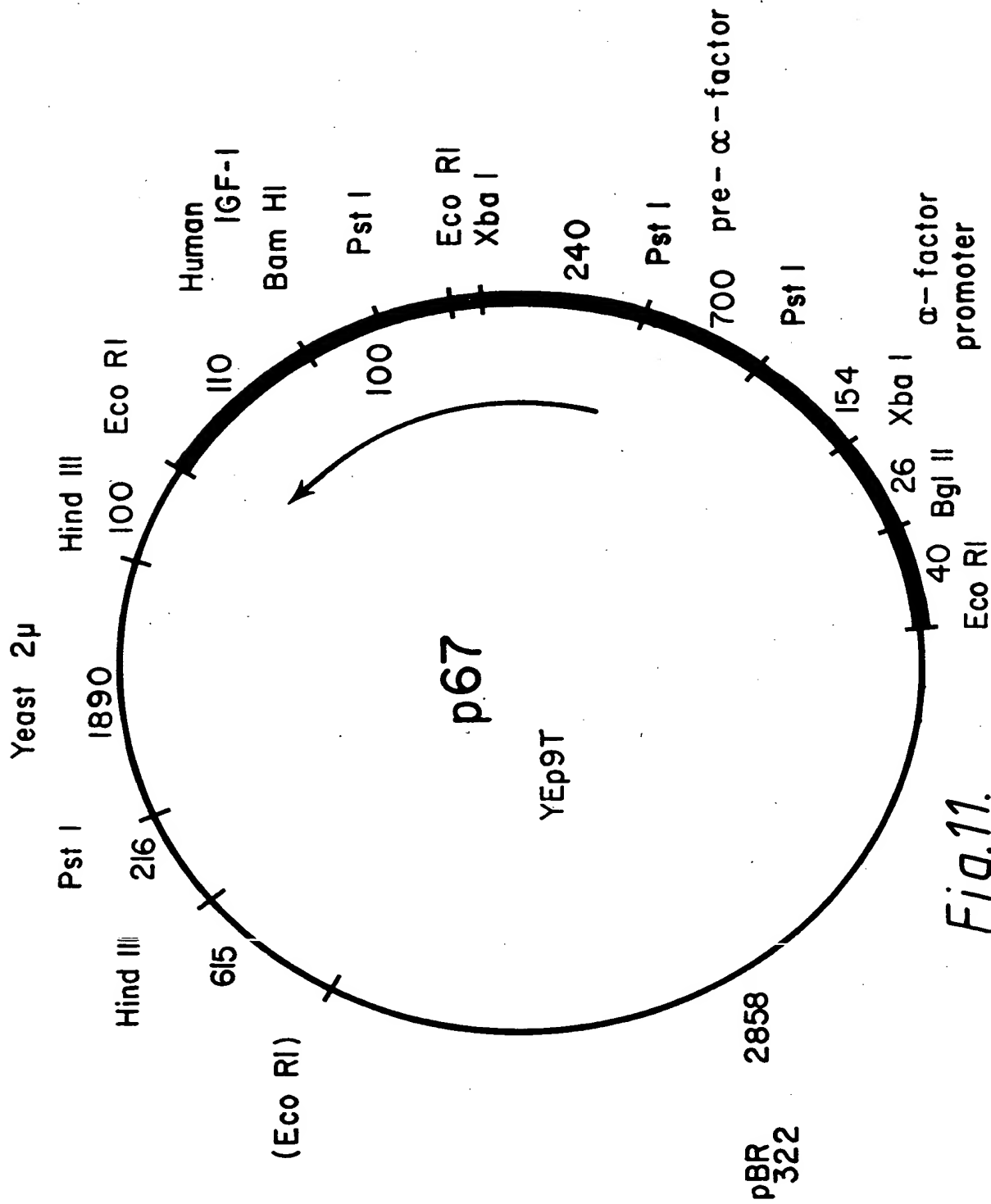


Fig.11.

Fig.12.

Yeast: Invertase Signal - IGF-1 Protein

```

1      met leu leu gln ala phe leu phe leu leu ala gly phe ala ala lys ile
    ATG CTT TTG CAA GCT TTC CTT TTC CTT TTG GCT GGT TTT GCA GCC AAA ATA

10

20      ser ala gly pro glu thr leu cys gly ala glu leu val asp ala leu gln phe val cys
    TCT GCA GGT CCG GAA ACT CTG TGC GGC GCT GAG CTG GTT GAC GCT CTG CAG TTC GTA TGT

30

40      gly asp arg gly phe tyr phe asn lys pro thr gly tyr gly ser ser arg arg ala
    GGT GAT CGA GGC TTC TAC TTC AAC AAA CCG ACT GGG TAC GGA TCC TCC TCT CGT CGT GCT

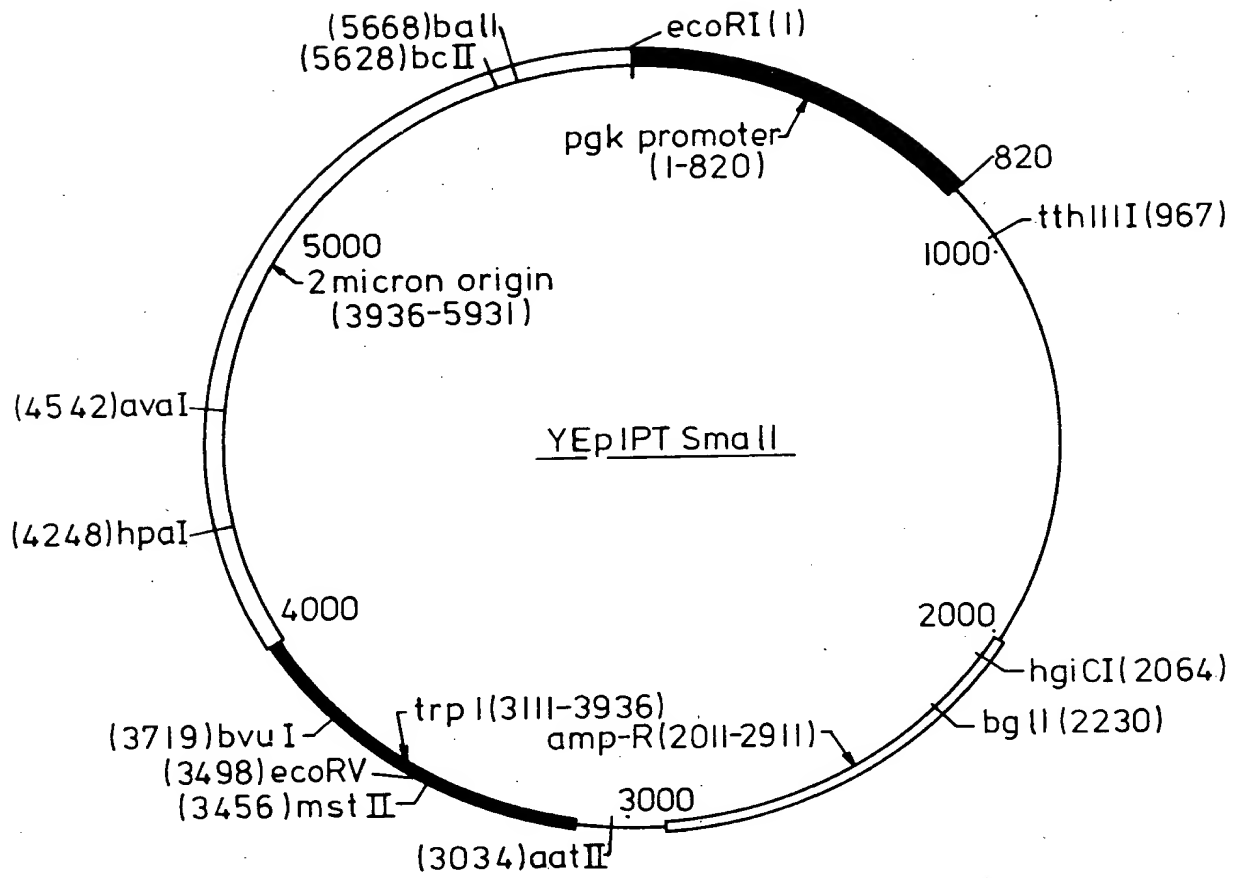
50

60      pro gln thr gly ile val asp glu cys cys phe arg ser cys asp leu arg arg leu glu
    CCG CAA ACC GGC ATC GTT GAT GAA TGC TGT TTT CGG TCC TGT GAC CTT CGC CGT CTG GAA

70

80      met tyr cys ala pro leu lys pro ala lys ser ala AM
    ATG TAC TGC GCT CCG CTG AAA CCG GCT AAG TCT GCA TAG
  
```

Underscored - Mature Human IGF-1

*Fig.13.*

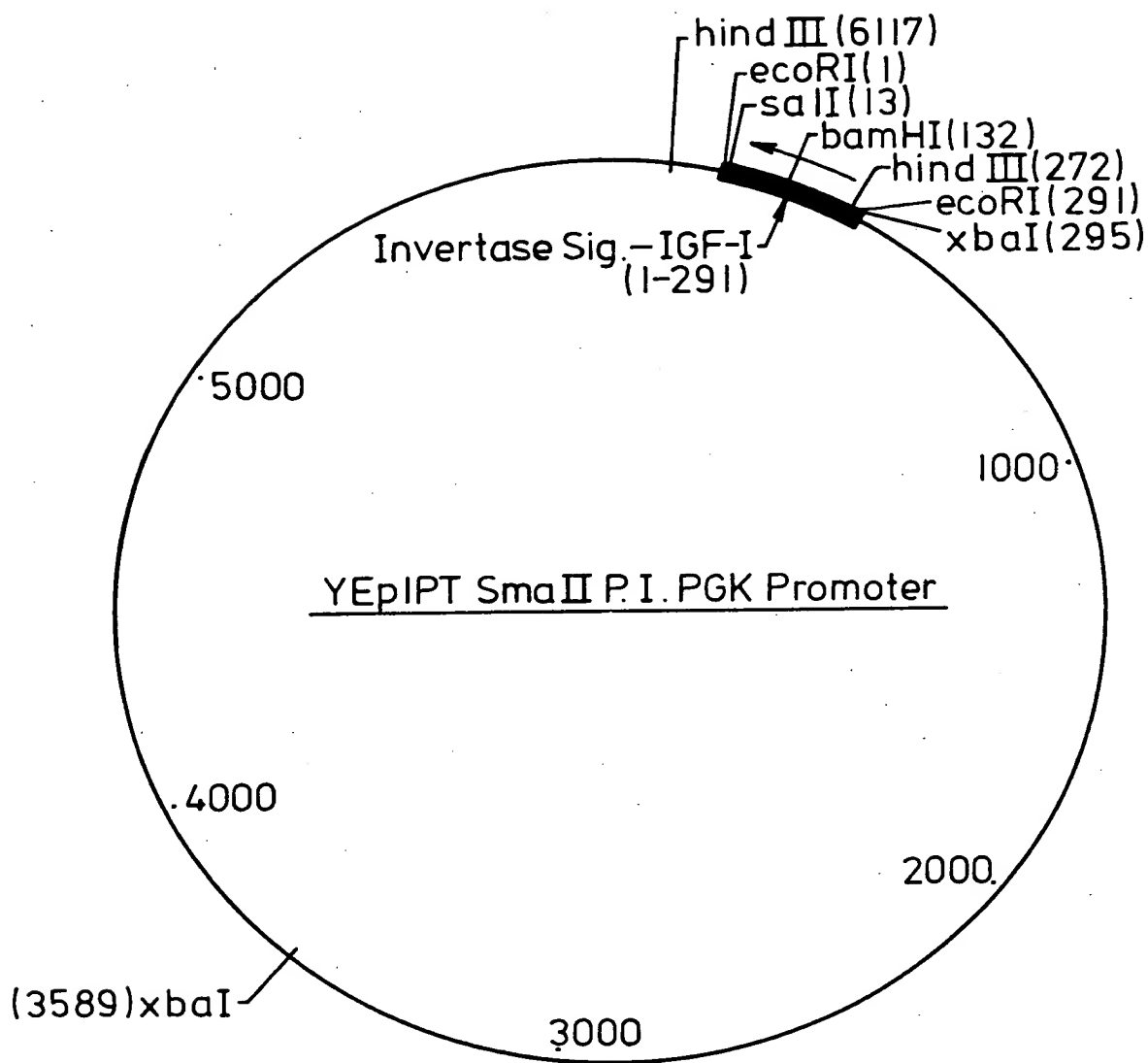
*Fig. 14.*

Fig. 15.

met arg phe pro ser ile phe ser ala val leu phe ala ala ser ser ala leu ala ala
ATG AGA TTT CCT TCA ATT TTT AGT GCA GTT TTA TTC GCA GCA TCC TCC GCA TTA GCT GCT

pro val asn thr thr thr glu asp glu thr ala gln ile pro ala glu ala val ile gly
CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA CAA ATT CCG GCT GAA GCT GTC ATC GGT

tyr ser asp leu glu phe gly asp phe asp val ala val leu pro phe ser asn ser thr asn
TAC TCA GAT TTA GAA GGG GAT TTC GAT GAT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT

asn gly leu leu phe ile asn thr thr thr ile ala ser ile ala ala lys glu glu gly val
AAC GGG TTA TTG TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA

ser leu asp lys arg asn ser asp ser glu cys pro leu ser his asp gly tyr cys leu
TCT TTG GAT AAA AGA AAC TCT GAC TCT GAA TGT CCA TTA TCG CAT GAT GGG TAC TGT TTG

his asp gly val cys met tyr ile glu ala leu asp lys tyr ala cys asn cys val val
CAC GAC GGA GTC TGT ATG TAT ATT GAA GCT CTA GAC AAG TAC GCT TGT AAC TGT GTT GTT

gly tyr ile gly glu arg cys gln tyr arg asp leu lys trp trp glu leu arg AM
GGT TAC ATC GGT GAA AGA TGT CAA TAC AGA GAT CTA AAG TGG TGG GAA TTG AGA TAG

Yeast Preinvertase - EGF

-19 met leu leu gln ala phe leu phe leu leu ala gly phe ala ala lys ile
 GAATTCATG ATG TTG TTG CAA GCT TTC TTG TTC TTG GCT GGT TTC GCT AAG ATC
 1 ser ala asn ser asp ser glu cys pro leu ser his asp gly tyr cys leu his asp gly
 TCT GCT AAC TCT GAC TCT GAA TGT CCA TTA TCG CAT GAT GGG TAC TGT TGT CAC GAC GGA
 10
 20 val cys met tyr ile glu ala leu asp lys tyr ala cys asn cys val val gly tyr ile
 GTC TGT ATG TAC ATT GAA GCT CTA GAC AAG TAC TAC GCT TGT AAC TGT GGT TAC ATC
 30
 40 gly glu arg cys gln tyr arg asp leu lys trp glu trp arg AM
 GGT GAA AGA TGT CAA TAC AGA GAT CTA AAG TGG TGG GAA TTG AGA TAG ATTGAATTGAATTGA
 50 53

AATCGATTAAAGCTT

*Fig. 18.*Human IGF-II coding sequence

1
 met ala tyr arg pro ser glu thr leu cys gly gly glu leu val asp thr
 ATG GCT TAT CGA CCA TCT GAA ACC TTG TGT GGT GGT GAG CTG GTG GAC ACC
 TCTAGAAATT

10
 20
 leu gln phe val cys gly asp arg gly phe tyr phe ser arg pro ala ser arg val ser
 CTC CAG TTC TGT GTC TGT GGG GAC CGC GGT TTC TAC TTC TCT AGG CCC GCA AGC CGT GTG AGC

30
 40
 arg arg ser arg gly ile val glu glu cys cys phe arg ser cys asp leu ala leu leu
 CGT CGC AGT CGT GGC ATC GTT GAG GAG TGC TGT TTC CGC AGC TGT GAC CTG GCC CTA TTG

50
 60
 glu thr tyr cys ala thr pro ala lys ser glu AM
 GAA ACC TAC TGT GCT ACC CCA GCT AAG TCT GAA TAG GATCC

68



European Patent
Office

EUROPEAN SEARCH REPORT

0128733

DOCUMENTS CONSIDERED TO BE RELEVANT			EP 84303783.9
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl. 7)
D, Y	<p>THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 253, no. 8, April 25, 1978, Baltimore, US</p> <p>E. RINDERKNECHT et al. "The Amino Acid Sequence of Human Insulin-like Growth Factor I and Its Structural Homology with Proinsulin" pages 2769-2776</p> <p>* Page 2769 (summary); fig. 1, 2 *</p> <p>--</p>	1	<p>C 12 N 15/00</p> <p>C 12 P 21/00</p> <p>A 61 K 37/36</p>
Y	<p>EP - A1 - 0 026 598 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA)</p> <p>* Claims 2, 14, 24, 30 *</p> <p>--</p>	1, 3, 5, 7, 9	
A	<p>EP - A2 - 0 022 242 (GENENTECH, INC.)</p> <p>* Claims 1, 5, 7, 10, 18 *</p> <p>--</p>	1, 3, 5, 7, 9	<p>TECHNICAL FIELDS SEARCHED (Int. Cl. 7)</p>
A	<p>THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 250, no. 11, June 10, 1975, Baltimore, US</p> <p>G. CARPENTER et al. "Characterization of the binding of ¹²⁵I-Labeled Epidermal Growth Factor to Human Fibroblasts" pages 4297-4304</p> <p>* Pages 4297 *</p> <p>----</p>	4	<p>C 12 N</p> <p>C 12 P</p> <p>A 61 K</p>
The present search report has been drawn up for all claims			
Place of search VIENNA		Date of completion of the search 06-09-1984	Examiner WOLF
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone</p> <p>Y : particularly relevant if combined with another document of the same category</p> <p>A : technological background</p> <p>O : non-written disclosure</p> <p>P : intermediate document</p> <p>T : theory or principle underlying the invention</p> <p>E : earlier patent document, but published on, or after the filing date</p> <p>D : document cited in the application</p> <p>L : document cited for other reasons</p> <p>& : member of the same patent family, corresponding document</p>			